



AN 94148842 EMBASE
DN 1994148842
TI Parenteral formulation development of renin inhibitor Abbott-72517.
AU Gupta S.L.; Patel J.P.; Jones D.L.; Partipilo R.W.
CS Clinical Center Pharmacy Department, NIH, Bldg 10, 9000 Rockville Pike, Bethesda, MD 20892, United States
SO PDA Journal of Pharmaceutical Science and Technology, (1994) 48/2 (86-91).
ISSN: 0279-7976 CODEN: JPHTEU
CY United States
DT Journal; Article
FS 027 Biophysics, Bioengineering and Medical Instrumentation
030 Pharmacology
037 Drug Literature Index
LA English
SL English
AB Abbott-72517 is an inhibitor of human renin and is being investigated for the treatment of hypertension. It is an orally bioavailable candidate which is being developed for oral as well as **intravenous** use. The preclinical development of this molecule involved studies to evaluate irritation at the site of injection in an animal model. Several formulation variables such as drug concentration, types of buffer (citrate or acetate), addition of **cosolvent (ethanol)** to enhance drug solubility, and tonicity modifiers such as glycerin or mannitol were evaluated. Additionally, in vitro formulation- whole blood hemolysis and plasma precipitation studies were conducted. Based on these studies, a liquid formulation containing 1.2 mg/mL Abbott-72517·HCl as base, 0.01M citrate buffer, pH 3.7, in 0.45% sodium chloride containing 2.5% mannitol was recommended for preclinical studies. Various processing and administration parameters were evaluated including filter qualification and compatibility of the drug with typical infusion fluids and administration sets. The liquid formulation was further characterized for physical and chemical stability. It was shown that it has acceptable stability at ambient temperature. Based on the accelerated temperature storage results, T90 at 25°C is > 1 year for the ready-to-use liquid formulation. Additionally, a lyophilized version of the liquid formulation was evaluated.

RESERVED. on STN

AN 94148842 EMBASE

DN 1994148842

TI Parenteral formulation development of renin inhibitor Abbott-72517.

AU Gupta S.L.; Patel J.P.; Jones D.L.; Partipilo R.W.

CS Clinical Center Pharmacy Department, NIH, Bldg 10, 9000 Rockville Pike, Bethesda, MD 20892, United States

SO PDA Journal of Pharmaceutical Science and Technology, (1994) 48/2 (86-91).
ISSN: 0279-7976 CODEN: JPTEU

CY United States

DT Journal; Article

FS 027 Biophysics, Bioengineering and Medical Instrumentation

030 Pharmacology

037 Drug Literature Index

LA English

SL English

AB Abbott-72517 is an inhibitor of human renin and is being investigated for the treatment of hypertension. It is an orally bioavailable candidate which is being developed for oral as well as **intravenous** use. The preclinical development of this molecule involved studies to evaluate irritation at the site of injection in an animal model. Several formulation variables such as drug concentration, types of buffer (citrate or acetate), addition of **cosolvent (ethanol)** to enhance drug solubility, and tonicity modifiers such as glycerin or mannitol were evaluated. Additionally, in vitro formulation- whole blood hemolysis and plasma precipitation studies were conducted. Based on these studies, a liquid formulation containing 1.2 mg/mL Abbott-72517·HCl as base, 0.01M citrate buffer, pH 3.7, in 0.45% sodium chloride containing 2.5% mannitol was recommended for preclinical studies. Various processing and administration parameters were evaluated including filter qualification and compatibility of the drug with typical infusion fluids and administration sets. The liquid formulation was further characterized for physical and chemical stability. It was shown that it has acceptable stability at ambient temperature. Based on the accelerated temperature storage results, T90 at 25°C is > 1 year for the ready-to-use liquid formulation. Additionally, a lyophilized version of the liquid formulation was evaluated.

ANSWER 116 OF 42399 EMBASE COPYRIGHT 2004 ELSEVIER INC. ALL RIGHTS

RESERVED. on STN

AN 94120900 EMBASE

DN 1994120900

TI Acute **intravenous** toxicity of dimethyl sulfoxide, polyethylene glycol 400, dimethylformamide, absolute **ethanol**, and benzyl alcohol in inbred mouse strains.

AU Montaguti P.; Melloni E.; Cavalletti E.

CS Via della Libertà km. 0.750, I-20052 Monza, Italy

SO Arzneimittel-Forschung/Drug Research, (1994) 44/4 (566-570).
ISSN: 0004-4172 CODEN: ARZNAD

CY Germany

DT Journal; Article

FS 052 Toxicology

LA English

SL English; German

AB Acute **intravenous** toxicity of some **solvents**, i.e. dimethyl sulfoxide (DMSO), polyethylene glycol 400 (PEG 400), dimethylformamide (DMF), absolute **ethanol** (EtOH) and benzyl alcohol (BeOH), was determined in three inbred (CD2F1, B6D2F1 and C57BL/6N) mouse strains used in many preclinical tests, mainly in oncology and toxicology. Haemolytic and precipitation potential tests in vitro were performed to assess the blood compatibility of the investigated **solvents** and its relationship with the observed symptoms. The single tested **solvents** did not show any major differences in acute toxicity in the three tested strains with the exclusion of DMSO (less toxic in CD2F1) and BeOH and EtOH (less toxic in B6D2F1). The tested dose ranges in the three strains (in ml/kg) were 1.0-5.66 for DMSO, 2.0-8.0 for PEG 400, 1.0-4.0 for DMF, 0.75-4.24 for EtOH, 0.025-0.4 for BeOH. The lowest tested dose was a safe dose and the highest one was the dose causing mortality in no more than half the animals in each group. The in vitro results suggest avoiding the use of BeOH (which also is more toxic than the other **solvents** in the in vivo test) and DMSO and using PEG 400, EtOH and DMF even though the latter induced a body weight decrease in the B6D2F1 mouse strain. As a general conclusion, dilution of these **solvents** in water is suggested to ameliorate their blood compatibility and the use of doses not higher than the lowest dose tested in this study is recommended.

ANSWER 116 OF 42399 EMBASE COPYRIGHT 2004 ELSEVIER INC. ALL RIGHTS RESERVED. on STN

AN 94120900 EMBASE

DN 1994120900

TI Acute **intravenous** toxicity of dimethyl sulfoxide, polyethylene glycol 400, dimethylformamide, absolute **ethanol**, and benzyl alcohol in inbred mouse strains.

AU Montaguti P.; Melloni E.; Cavalletti E.

CS Via della Liberta km. 0.750, I-20052 Monza, Italy

SO Arzneimittel-Forschung/Drug Research, (1994) 44/4 (566-570).
ISSN: 0004-4172 CODEN: ARZNAD

CY Germany

DT Journal; Article

FS 052 Toxicology

LA English

SL English; German

AB Acute **intravenous** toxicity of some **solvents**, i.e. dimethyl sulfoxide (DMSO), polyethylene glycol 400 (PEG 400), dimethylformamide (DMF), absolute **ethanol** (EtOH) and benzyl alcohol (BeOH), was determined in three inbred (CD2F1, B6D2F1 and C57BL/6N) mouse strains used in many preclinical tests, mainly in oncology and toxicology. Haemolytic and precipitation potential tests in vitro were performed to assess the blood compatibility of the investigated **solvents** and its relationship with the observed symptoms. The single tested **solvents** did not show any major differences in acute toxicity in the three tested strains with the exclusion of DMSO (less toxic in CD2F1) and BeOH and EtOH (less toxic in B6D2F1). The tested dose ranges in the three strains (in ml/kg) were 1.0-5.66 for DMSO, 2.0-8.0 for PEG 400, 1.0-4.0 for DMF, 0.75-4.24 for EtOH, 0.025-0.4 for BeOH. The lowest tested dose was a safe dose and the highest one was the dose causing mortality in no more than half the animals in each group. The in vitro results suggest avoiding the use of BeOH (which also is more toxic than the other **solvents** in the in vivo test) and DMSO and using PEG 400, EtOH and DMF even though the latter induced a body weight decrease in the B6D2F1 mouse strain. As a general conclusion, dilution of these **solvents** in water is suggested to ameliorate their blood compatibility and the use of doses not higher than the lowest dose tested in this study is recommended.


```

Seq primer: M13-21
Class: BAC ends.
FEATURES
    source
        1..337
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="2534J20"
            /clone_lib="CITBI-E1"
            /sex="male"
            /cell_type="sperm"
            /note="Vector: pBelorAC11; Site_1: EcoRI; Site_2: EcoRI;
            Caltech Human BAC Library D"
BASE COUNT      78 a   70 c   75 g   114 t
ORIGIN
    Query Match      32.8%; Score 328; DB 17; Length 337;
    Best Local Similarity 98.5%; Pred. No. 2.9e-32;
    Matches 331; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 353 A A A T T A G C C A G G T G T G T G C A C A G C C T G A G T C C C A G C T A C T T G G G A G G C T G A G G G 412
Db 337 A G A A T T A G C A G G T G T G T G C A C A G C C C T G A G T C C C A G C T A C T T G G G A G G C T G A G A G 278
QY 413 G A G A A T C C G T T G A C C C A G T A G C A G G T T G C A G T G A G C C G A G A T A A G A G T C A C T G C A 472
Db 277 C A G A A T C C G T T G A C C C A G T A G C A G G T T G C A G T G A G C C G A G A T A A G A G T C A C T G C A 218
QY 473 C T C C A G C C T G G T G A C A G A G A G A C T C C C T C T C A G A A A T A A A A T A A A A T A A A A 532
Db 217 C T C C A G C C T G G T G A C A G A G A G A C T C C C T C T C A G A A A T A A A A T A A A A T A A A A 158
QY 533 T A A A T A A A T A A A T A A A T A A A T T C T A A A G G G C T G C C A T T T G C T A G C A C T A T A T 592
Db 157 T A A A T A A A T A A A T A A A T A A A T T C T A A A G G G C T G C C A T T T G C T A G C A C T A T A T 98
QY 593 G C C C A A T A A G T A G C T A C A T A T C C C A C C C T A C C A G T G C T G A A A T T A G T T C 652
Db 97 G C C C A A T A A G T A G C T A C A T A T C C C A C C C T A C C A G T G C T G A A A T T A G T T C 38
QY 653 T T T T G T C A C C C C C C A T T A G A C T T A A G C A G A A T T C 688
Db 37 T T T T G T C A C C C C C C A T T A G A C T T A A G C A G A A T T C 2

RESULT 2
AG155188/c
LOCUS      AG155188      667 bp      DNA      linear      GSS 09-JAN-2002
DEFINITION Pan troglodytes DNA, clone: RP43-019G22.TJ, genomic survey
ACCESSION AG155188
VERSION    AG155188.1 GI:16684866
KEYWORDS  GSS.
SOURCE     Pan troglodytes
ORGANISM   Pan troglodytes male lymphocytes DNA, clone_lib:RPCI-43 Chimpanzee
REFERENCE  1
AUTHORS    Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
            Totoki, Y., Watanabe, H. and Sakaki, Y.
            BAC end sequences of Library RPCI-43
            Unpublished
            2 (bases 1 to 667)
            Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
            Totoki, Y., Watanabe, H. and Sakaki, Y.
            Direct Submission
            Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan
            (E-mail: chimpanzee@sc.riken.go.jp URL: http://hgp.gsc.riken.go.jp/
            Tel: 81-45-503-9111, Fax: 81-45-503-9170)
            Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
PRIMERS
LIBRARY
    Sequencing: TJ
    Vector : pBACe3.6
    R.Site 1 : EcoRI
    R.Site 2 : EcoRI
    Location/Qualifiers
        1..667
            /organism="Pan troglodytes"
            /db_xref="taxon:9598"
            /clone="RP43-019G22.TJ"
            /sex="male"
            /cell_type="lymphocytes"
            /clone_lib="RPCI-43 Chimpanzee Male BAC Library"
BASE COUNT      140 a   167 c   150 g   210 t
ORIGIN
    Query Match      30.7%; Score 306.8; DB 17; Length 667;
    Best Local Similarity 80.1%; Pred. No. 7.9e-93;
    Matches 411; Conservative 0; Mismatches 93; Indels 9; Gaps 4;
QY 66 T G T A A A A A A A T T A A T T G A A T A G C T A G C G C G G T G C T C A C G C T G T A A T C C C A G A C T 125
Db 515 T C T A C T A A A A A T A C A A A A A T T G T T G G C G T A G T T G G C G C T G T G T A A T C C C A G A C T 456
QY 126 T T A G A G S T C G A G G G T G A T C A C T T G A G G T C A G G A G T T T T G A G A C C A G C C T G G C C A A 185
Db 455 T T G G G G C C C A G G G G G T G A T C A C T G A G G T C A G G A G - T T C A A G A C C A G C C T G G C C A A 397
QY 186 C A C G S T G A A A C C C A C T C T A C T A A A A A T A A A A A T T A G C T T N G G G T G C G S T G G C T C A C A C 245
Db 396 C A T G G C A A A C C T C C C C T C T A C T A A A A T A C A A A T T A G C C A G C A T G T G G C T C A C A C 337
QY 246 C T G T A T C C C A G C A C T T T G G G A G G C T G A G A C G S T G A G A C C C T G A G T C A G G A G T T C A A 305
Db 336 C T G T A T C C C A G C A C T T T G G G A G G C T G A T G C G S T G A T C A C C T G A G G T C A G G A G T T C A A 277
QY 306 G G C C A C C T G G C C A C A T G T G A A C C A G C G T C T C T A T A A A A T A C A A A A T T A G C C A G G 365
Db 276 G A C C A C C T G G C C A C A T G T G A A C C C A C T C T C T A T A A A A T A C A A A A T T G G T G G 217
QY 366 T G T G - G T G C A C A G C C C T G A G T C C C A C T A C T T G G G A G C T G A G C G G A G A T C G C T T 424
Db 216 I C T T A G T G C A G G T G C C T G T A T C A T G C T A C T T G G A G G C T G A G G C A G G A A T C A C T G 157
QY 425 G A C C C A G T A G C A G A G T T G C A G T A G C C G A G A T A A G A G T C A C T G C A C T C C A G C C T G G G 484
Db 156 G A C C C A G G C A G A G G T T G C A G T G A G A T C A - T G C C A C T G C A C T C C A T C C T G G G 98
QY 485 T G A C A G A C A A G A C T C C C T C T C A G A A A T A A A T A A A T A A A T A A A T A A A T A A A T A A A T A 544
Db 97 T G A T A G C A G A C T C T G C T C A A A A A T A A A A T A A A A T A A A A T A A A A T A A A A T A A A T A A 44
QY 545 A A A T A A A T A A A T T C T A A A G G G T G C A T T T 577
Db 43 G A A T A A A T A A A T G A G A A A C C C G G G T C T T T 11

RESULT 3
AG011265
LOCUS      AG011265      713 bp      DNA      linear      GSS 07-FEB-1999
DEFINITION Homo sapiens genomic DNA, 21q region, clone: S594SP2Nh, genomic survey sequence.
ACCESSION AG011265
VERSION    AG011265.1 GI:3370908
KEYWORDS  GSS.
SOURCE     Homo sapiens DNA, clone: S594SP2Nh.
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 713)
AUTHORS    Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.

```

TITLE Homo sapiens genomic DNA, chromosome 21q
REFERENCE Published Only in DataBase (1998)
AUTHORS 2 (bases 1 to 713)
TITLE Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
JOURNAL Direct Submission
FEATURES Submitted (30-JUL-1998) Masahira Hattori, Kitasato University,
Department of Science, JST Sequencing Laboratory, Kitasato 1-15-1,
Sagamihara 228, Japan (E-mail:hattori@cnc.ims.u-tokyo.ac.jp,
Tel:0427-78-9732, Fax:0427-78-9561)
SOURCE Location/Qualifiers
1 713

```

4. /13
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q"
/clone="5594Sp2Nh"
168 a 174 c 175 g 156 t 40 others
BASE COUNT
ORIGIN

```

Query Match 29.5%; Score 295; DB 17; Length 713;
Best Local Similarity 78.7%; Pred. No. 2.2e-28;
Matches 363. Conserved Regions 1.

[illegible]

RESULT 4	
BC018568/c	
LOCUS	3001 bp
DEFINITION	Homo sapiens, clone IMAGE:3446135, mRNA.
ACCESSION	BC018568
VERSION	BC018568.1 GI:17235741
KEYWORDS	HTC.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens.
	linear HTC 03-DEC-2001

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 3001)
Strausberg R.
Direct Submission
Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590.

REMARK	COMMENT

USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cdapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalon@bcm.tmc.edu
Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X.,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu,
Munzy, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 19 Row: k Column: 5
 This clone has the following problem: no 5' EST match.

FEATURES source

```

source
1 300H/300dalliers
/organism="Homo sapiens"
/bi_xref="taxon:9606"
/clone IMAGE:3446135
/tissue_type="Placenta, cho
/clone_lib= NIH_MGC_10
/lab_host=DHI08
/notes=Vector: pcwv-sport6"
715 a 688 c 693 g 905 t
BASE COUNT
ORIGIN

```

Query Match 29.4%; Score 294.2; DB 11; Length 3001;
Best Local Similarity 80.5%; Pred. No. 9.9e-29;
Matches 300

QY	52	ACGTGCTAGATTGTTGTAAAAAATTAATAGAAATAGCTAGCGCGGTGCTCAGCCCT	111
Db			
QY	474	ATCTGCCAAGAAATTTAAAGCAGAGTCATAAGACACCGCTGGCGGGTGGCTCAGCCCT	415
Db			
QY	112	GTAATCCCGACACTTTAGAGGGTGAAGAGGGTGATGACTTGAGTGCAGGAGTTTGTAG	171
Db			
QY	414	GTATCCCGACACTTTGTAGCGCCAGAGCGGCGCATCACTGAGTGCAGGAG-TTTGAA	356
Db			
QY	172	ACCAGCGTGGCCACAACGGTGAACCCCATCTCTACTAAAAATAAAAAATAGCTNGSGT	231
Db			
QY	355	ACCAGCGTGGACAACATGGTGAACCCCATCTCTACTAAAAATAGAAAAATTACCGAGGT	296
Db			
QY	232	CGGCTGGCTCACACCTGTATCCACGACTTTGGAGGCGGTGAGACGGGTGATCACTGA	291
Db			
QY	295	GCAGTGGCTTACGCGTGTATCCCGACACTTTGTAGCGCGGAGCGGCAGATCACTGA	236
Db			
QY	292	AGTCAGGAGTTCAAGGCCGCGCTGGGCAACATGCTGAACACGCTCTCTACTAAAAATAC	351
Db			
QY	235	GGTCAGGAGTTTGAACACCGCTGGACAACATGCTGAACCCCTCTCTACTAAAAATAC	176
Db			
QY	352	AAAAATTAGCAGAGTGTGTGGCACACGGCTGTAGTCCCGTACT-TGGGAGCGTGAAG	410
Db			
QY	175	AAAAATAGCTGGGTTGGTGGTGGGTACCTGTGTAGTCCCGTACTCTGGGAGGCTGAGG	116
Db			
QY	411	CGGAAGATTCGCTTGAACCCACTTAGCCAGAGTTGCAGCTGAGCCGAGATAGAGTCACTG	470
Db			
QY	115	CACGAGAATTGCCTTGAACCCCGAGGCTGGAGTTGTAGTGACCCATATCA-TGCCACCG	57
Db			
QY	471	CATCTCCAGCGTGGGTGACAGACCAAGACTCCCTCTCAGAAAAATAAAAAATAAA	522
Db			
QY	56	CACTCCGCTTGGGTGACAGACGAGCTCCATCTCAAAAAAATAAAAAAATAAA	5
Db			

RESULT 5	
BQ710535	
LOCUS	
DEFINITION	

LOCUS	BQ710535	986 bp	linear	EST 16-JUL-2002
DEFINITION	AGENCOURT_8352291	NIH_MGC_113	Homo sapiens	cDNA clone IMAGE:6277543

5', mRNA sequence.
 BQ710535
 VERSION BQ710535.1 GI:21849434
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Mammalia; Euteleostomi;
 1 (bases 1 to 986)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 UNpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@email.nih.gov
 Tissue Procurement: Dr. Mark Watson
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Cloning Distribution by: Agencourt Bioscience Corporation
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L12M2464 row: b column: 08
 High quality sequence set: 699.
 Location/Qualifiers
 1..986
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6277543"
 /clone_lib="NIH_MGC_113"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: spleen; Vector: pOTB7; Site:1: XhoI; Site:2:
 EcoRI; cDNA made by oligo-dT priming. Directionally clone
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."
 BASE COUNT 284 a 231 c 243 g 225 t 3 others
 ORIGIN
 Query Match 29.2%; Score 292; DB 14; Length 986;
 Best Local Similarity 82.0%; Pred. No. 4.2e-28;
 Matches 396; Conservative 0; Mismatches 81; Indels 6; Gaps 5
 QY 54 CTGCTAGATTGTTGTAATAAATTAATGAATAGCTAGGCGGCTGCTACGCTCT 113
 Db 205 CTTCGCTTAGTTTCTTAAGCAAGATTATTAAGCCAGGCAAGCTGGCCAGCTGT 264
 QY 114 AATCCAGCACTTTAGAAAGTGTGAAGAGGGTGGATCATTGAGTCTAGGATTTGAGAC 173
 Db 265 AATCCAGCACTTTGGAGGCGAGGTGGGTGGATCACCCTGAGTCTAGGAGTCAAGAC 323
 QY 174 CAGCTGGCCACACAGGTGAACCCCACTCTACTAAAAATA-AAAAATAGCTGGGTG 232
 Db 324 CAGCTGGCCACACATGTTGAACCCCGCTCTACTAAAAATAGAAAAATTAGCTGTGTA 393
 QY 233 CGGTGCTCACAACCTGTAATCCAGCACTTTGGAGGCTGAGAGCGGTGGATCACTGAA 292
 Db 384 CGGTGCTCA-GCTTGTATCCAGCACTTTGGAGGCTGAGGCGGTGGATCAG-CAG 440
 QY 293 GTCAGAGTTCAAGGCCAGCCCTGGGCAACATGGTGAACACCACTCTCTACTAAAAATACA 352
 Db 441 GTCAGAGATCGAGACTGTCTTGGTCAACATGGTGAACCCCATCTCTACTAAAAATACA 500
 QY 353 AAAATTAGCAGGTGTGGTGGCAACCCCTGTAGTCCAGCTACTTTGGAGGCTGAGCG 412
 Db 501 AAAATTAACCTGGGTGTAGTGGCAACCCCTGTAGTCCAGCTACTTTGGAGGCTGAAGCA 560
 QY 413 GAAGATCGCTTGAACCCAGTAGAGGCAAGGTTGCAGTGCAGCCAGATAGAGTCACTGCA 472
 Db 561 GGAGATCGCTTGAACCCGAGGCGAGTGTGCAGTGCAGCCAGATAGAGTCACTGCA 619

```

|||||
Db 441 CCTGACCAACATGGTGAACCCCTGCTCTACTAAAAATACAAAAATTAGCGGTGTGGT 500
QY 372 GGCACAGCCTGTAGTCCCACTACTTGGAGGCTGAGGGGGAAGATCGCTTGAACCCA 431
Db 501 GGTGGCGGCTCTAATCCAGCTACTCGGAGGCTGAGGCGGAGAAATTCCTTGAACGG 560
QY 432 GTAGGAGAGTTCAGTGCAGCCGATGAAGTCACTGCACTCCAGCCTGGGTGACAGA 491
Db 561 GGAGCGGAGGTTCAGTGAACAGGAT-CGCGCACTGCACTCCAGCCTGGCGACAGA 619
QY 492 GCAAGACTCCCTCTCAGAAAAATAAAA 517
Db 620 GCGAGACTCGCTCTCAAAAAAATAAAA 645

RESULT 7
A0514325 543 bp DNA linear GSS 05-MAY-1999
LOCUS HS_5194.B2.H07.T7A.RPCI-11 Human Male BAC Library Homo sapiens
DEFINITION genomic clone Plate-770 Col=14 Row=P, DNA sequence.
ACCESSION A0514325
VERSION A0514325.1 GI:4746616
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 543)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pierder@u.washington.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Research Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 770 row: P column: 14
Seq primer: T7
Class: BAC ends
High quality sequence stop: 543.
FEATURES
source
Location/Qualifiers
1..543
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=770 Col=14 Row=P"
/clone_lib="RPCI-11 Human Male BAC Library"
/sex="male"
/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI methylase. Size selected DNA was cloned into the
pBAC3.6 vector at EcoRI sites"
BASE COUNT 145 a 123 c 152 g 113 t 10 others
ORIGIN

Query Match 29.0%; Score 290; DB 17; Length 543;
Best Local Similarity 77.6%; Pred. No. 1.2e-27;
Matches 384; Conservative 0; Mismatches 107; Indels 4; Gaps 3;
QY 40 CGTAAATGAAGAAGCTCTAGATTGCTTAAATAAATTAATGGAAATAGGCTAGCGGG 99

```

```

|||||
Db 22 CTTGNATGTATGCANTTCGGGAATTCCTGATAGATTNTANAGTGGAGCCAGGTGCGG 81
QY 100 TGGCTCAGCGCTGTAAATCCAGCACTTTAGAGGTCGAAGAGGTTGGATCATTGAGGTC 159
Db 82 TGGCTCAGCACTGTAAATCCAGCACTTTAGAGGTCGAAGAGGTTGGATCATTGAGGTC 141
QY 160 AGGATTTTGGAGCCAGCCTGGCCAAACACAGTGAACCCCACTCTACTTAAATAAATA 219
Db 142 AGGAG-TTGGAGAGCAGCCTGGCCAAACATGTTGAACCTCTCTCTACTTAAATAAATA 200
QY 220 ATTAGCTNGGTCGGTGGCTCAGCACTTGAATCCAGCACTTTGGAGGCTGAGACGGG 279
Db 201 AAATTCGAGGCAATGATGGCTATGCTGTAATCCAGCGCTTTGGAGGCTGAGCANG 260
QY 280 TGGATCAGCTGAAGTCAGGAGTTCAGGCGGAGCCTGGGCAACATGGTGAACACAGCTC 339
Db 261 AGGATCAC--GAGTCACTAGTTTGACAGCAGCCTGGGCAATATGGTGAACCCCGTCTC 318
QY 340 TACTTAAATAAATAAATAATAGCAGGTCGGTGGCAGCAGCCTGTAGTCCAGCTACTTG 399
Db 319 TACTTAAATAAATAAATAATAGCAGGTCGGTGGCAGCAGCCTGTAGTCCAGCTACTTG 378
QY 400 GGAGGCTGAGGCGGAGATCGCTTGAACCCAGTAGGAGGTCAGTGGAGCCGAGAT 459
Db 379 GGAGGCTGAGGCGGAGATCGCTTGAACCCAGTAGGAGGTCAGTGGAGCCGAGAT 438
QY 460 AAGAGTCACTGACCTCCAGCCTGGGTGACAGAGCAAGACTCCCTCTCAGAAATAAATA 519
Db 439 -CGTGCCACTGCTCCAGCCTGGGACAGAGCAAGACTCTGTCTCTCATTAACAAACAAC 497
QY 520 AATATAAATAAATA 534
Db 498 AAACAAGAAAAATA 512

RESULT 8
A0426757 542 bp DNA linear GSS 24-MAR-1999
LOCUS CITBI-EI-2572B9.TF CITBI-EI Homo sapiens genomic clone 2572B9, DNA
DEFINITION sequence.
ACCESSION A0426757
VERSION A0426757.1 GI:4499437
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 542)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
Map Building
JOURNAL Unpublished (1997)
COMMENT Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbest@igrr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends
Location/Qualifiers
1..542
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2572B9"
/clone_lib="CITBI-EI"
/sex="male"

```



```

RESULT 12
BG110480      545 bp      mRNA      linear      EST 30-JAN-2001
602278903F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4366465 5',
mRNA sequence.
BG110480
BG110480.1 GI:12503986
EST.
human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 545)
NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS      National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE        Unpublished (1999)
JOURNAL
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
cDNA distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10017 row: f column: 02
High quality sequence stop: 545.
FEATURES
Location/Qualifiers
1..345
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4366465"
/clone.lib="NIH_MGC_86"
/tissue_type="osteosarcoma, cell line"
/lab_host="DH10B (phage-resistant)"
/site="organ: bone; Vector: pCMV-SPORT6; Site.1: NotI;
Site.2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.533 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT    159 a 122 c 135 g 119 t
ORIGIN
Query Match      28.6%; Score 285.8; DB 12; Length 545;
Best Local Similarity 81.1%; Pred. No. 3.9e-27;
Matches 368; Conservative 0; Mismatches 83; Indels 3; Gaps 3;
QY 79 AATGGATAGGTAGCGCGGTGCTCAGCGCTGTAAATCCAGCACCTTTAGAAGGTGAA 138
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 93 ATGCTCTAGTGCAGCGGTGGTGGTGCATCTAATCCAGCACTTTGGAGGCGGAG 152
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 139 GAGGGTGGATCACTTGAGTGCAGAGTTTGGAGACCGCTGGCCACACCGTGAACACC 198
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 153 GCATGTGGATCACTGTATCAGG-GTTTGAACACCGCTGGCCAAATGTTGAACACC 211
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 199 CATCTCTACTAAAAATAAAAAATTAGCTTNGGGTGGGTGGCTCACACCTGTAAATCCAGC 258
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 212 CATCTCTACTAAAAATACAAAATGTCGCCAGGTGCGATTTGCTCATGCTGTATTCACAGC 271
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 259 ACTTTGGAGGCTGAGACGGGTGGATCACCCTGAAGTCAGGAGTTCAAGGCCAGCGCTGGGC 318
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 272 ACTTTGAGACACCGAGGAGGCGAGATCACCCTGAGGCTCAGGAGTTTGAGACCAGCGCTGGCC 331
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 319 AACATGGTGAACACCACTCTCTACTAAAAATACAAAATTAGCCAGGTGTGGTGGCACAC 378
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 332 AACATGGTGAACACCTATCTCTACTAAAGTACAAAATTAGCCTGGTGTGGTGGCAGCGC 391
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 379 GCCTGTAGTCCACGCTACTTTGGAGCGCTGAGCGGAGAAATCGCTTGAACCCAGTAGCA 438
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 392 ACTGTGAATTCAGCTACTCAGGAGCGCTGAGCGAGGAGAAATCACTTGAACCCAGGAGCGC 451
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 439 GAGGTTGCAGTGAGCCGAGATAGAGTCACTGTGACCTCCAGCTGGGTGCAC-AGAGCAAGA 497

```

```

QY 300 GTTCAAGCCAGCGTGGCAACATGGTGAACACCTCTCTACTATAAATAACAAAAATTA 359
    ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 207 GTTCCAGACCGCTGACCAATATGGTGAACCCCATCTCTACTAAAAACACAAAAATTA 148
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 360 GCCAGGTGTGGTGGCACAGCGCTGTAGTCCAGCTACTTGGAGGCTGAGGCGGAGAAAT 419
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 147 GCCAGGTGTGGTGGCGGCGCTGTAGTCCAGCTACTTGGAGGCTGAGGCGAGAGAAAT 88
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 420 CCCTTTGAACCCACTAGGACAGAGGTTCAGTGGAGCGCGAGATAGAGTCACTGCATCTCCAGC 479
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 87 CCCTTTGAACCCGGGAGATGGAGGTTCAGTGGAGCGCGAGATCA-CGCCACTGCATCTCCGC 29
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 480 CTGGGTGACAGACAGAGACTCCCTCTCA 507
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 28 CTGGACGACAGAGGAGACTGTATCTCA 1

RESULT 14
AQ484455/c
LOCUS
DEFINITION
    , DNA sequence.
    , 568 bp DNA linear GSS 24-APR-1999
    , RPCI-11-243K18.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-243K18
ACCESSION
    AQ484455
VERSION
    AQ484455.1 GI:4671776
KEYWORDS
    GSS.
SOURCE
    human.
ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
    1 (bases 1 to 568)
    Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter
    , J.C.
    Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
    Map Building
    J. Biol. Chem. 272:10000-10004 (1997)
    Unpublished (1997)
    Other_GSSs: RPCI-11-243K18.TV
    Contact: Shaying Zhao, William Nierman, Mark Adams
    Department of Eukaryotic Genomics
    The Institute for Genomic Research
    9712 Medical Center Dr., Rockville, MD 20850
    Tel: 301 838 0200
    Fax: 301 838 0208
    Email: hbeetigr.org
    Clones are derived from the human BAC library RPCI-11. For BAC
    library availability, please contact Pieter de Jong
    (pieter@dejong.med.buhalo.edu). Clones may be purchased from
    BACPAC Resources (http://bacpac.med.buhalo.edu/ordering) or from
    Research Genet cs (info@resgen.com). BAC end search page:
    http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
    Seq primer: SP6
    Class: BAC ends.
FEATURES
    source
        1..568
        /organism="Homo sapiens"
        /db_xref="GDB:7593185"
        /db_xref="taxon:9606"
        /clone="RPCI-11-243K18"
        /clone_lib="RPCI-11"
        /sex="Male"
        /cell_type="Lymphocytes"
        /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
        RPCI11 Human Male BAC Library"
BASE COUNT
    112 a 152 c 143 g 161 t
ORIGIN

Query Match
    Best Local Similarity 28.5%; Score 284.8; DB 17; Length 568;
    Matches 359; Conservative 0; Mismatches 88; Indels 2; Gaps 2;
QY 78 AAATGGAATAGCTAGCGCGGTGGCTGACGCTGTATCCAGCACTTTAGAAAGTCA 137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 540 AAAATGATCGTCCGGGTGGTGGCTCACCTGTATCCAGCACTATGTTAGTACCA 481
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 138 AGAGGTGGATCACTTGGAGTTCAGGAGTTTTCAGACACAGCTGGCCACACAGGTGAACC 197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 480 GGCAGGTGGATCACTTGAAGTTCAGGAG-TTTGAGACACAGCTGGCCACACATGGTGAACC 422
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 198 CCATCTCTACTATAAATAAATAAATAGCTGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 257
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 CCATCTCTACTATAAATAAATAAATAGCTGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 362
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 258 CACTTTGGAGGCTGAGACGGGTGGATCAGCTGAAGTTCAGGAGTTTCAAGGCCAGCTGGG 317
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 CACTTTGGAGGCTGAGACGGGTGGATCAGCTGAAGTTCAGGAGTTTCAAGGCCAGCTGGG 302
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 318 CAACATGGTGAACACAGCTCTCTACTATAAATAAATAAATAGCTGCGGTGGTGGTGGTGGT 377
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 CAACATGGTGAACACAGCTCTCTACTATAAATAAATAAATAGCTGCGGTGGTGGTGGTGGT 322
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 378 CGCTGTAGTCCAGCTACTTGGAGGCTGAGGCGGGAAGTTCGCTTGAACCCAGTGGC 437
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 TGCTGTATCCAGCTACTTGGAGGCTGAGGCGGGAAGTTCGCTTGAACCCAGTGGC 182
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 438 AGAGTTCAGTGGAGGCTGAGGCGGGAAGTTCGCTTGAACCCAGTGGC 497
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 GGAGGCTGAGTGGAGGCGGGAAGTTCGCTTGAACCCAGTGGC 123
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 498 CTCCTCTCTCAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 526
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122 CTCCTCTCTCAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 94
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
BQ437946
LOCUS
DEFINITION
    BQ437946
    5', mRNA sequence.
ACCESSION
    BQ437946
VERSION
    BQ437946.1 GI:21177022
KEYWORDS
    EST.
SOURCE
    human.
ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
    1 (bases 1 to 895)
    NIH-MGC http://mgi.nci.nih.gov/.
    National Institutes of Health, Mammalian Gene Collection (MGC)
    Unpublished (1999)
    Contact: Robert Strausberg, Ph.D.
    Email: cgapbs@mail.nih.gov
    Tissue Procurement: ATCC/DCTD/DTF
    CDNA Library Preparation: Life Technologies, Inc.
    DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
    Clone distribution by: Agencourt Bioscience Corporation
    found through the I.M.A.G.E. Consortium/LLNL at:
    http://image.llnl.gov
    Plate: LLAM13503 row: o column: 20
    High quality sequence stop: 641.
FEATURES
    source
        1..895
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:6157675"
        /clone_lib="NIH-MGC_72"
        /tissue_type="melanotic melanoma"
        /lab_host="DH10B (phage-resistant)"
        /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
        Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
        Average insert size 2 Kb. Library constructed by Life
        Technologies."
BASE COUNT
    246 a 200 c 258 g 190 t 1 others
ORIGIN

Query Match
    28.4%; Score 284; DB 14; Length 895;

```

Best Local Similarity 80.5%; Pred. No. 4.5e-27;
Matches 381; Conservative 0; Mismatches 86; Indels 6; Gaps 4;

```

QY 98 GGTGGCTCAGCCCTGTATCCAGCACTTTAGAAGTCTGAAGAGGGTGGATCAGCTTGAGG 157
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3 GGTGGCTGAGCCCTGTATCCAGCACTTTGGGAAGCTGAGGAGGGTGGATCAGCTTGAGG 62
QY 158 TCAGGAGTTTTCAGACCGCTGGCCACACAGCTGAACCCCATCTCTACTAAAAATAA 217
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
63 TCAGGAG-TTCAGACCGAGCTGGCCACACATGGTTAAACCCCTGTCTACTAAAAATACA 121
QY 218 AATTTAGCTNNGGTCGGTGGCTCAGACCTGTATCCAGCACTTTGGGAGGCTGAGACG 277
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
122 AAAATTGGCGGGCGCAGTGGCTCACTCTGTATCCAGCACTTTGGGAGGCGGCGG 181
QY 278 GGTGATCAGCTGAAGTCAGGAGTTCAGGCCAGCCTGGGCAACATGGTGAACCCAGCTC 337
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
182 GCGGATCAC--GAGGTCAGGAGATTGAGACCATCTCTGGCTAACACGGTGAACCCGCTC 239
QY 338 TCTACTAAAAATACAAAA--ATTAGCCAGGTGTGGTGGCACAGCGCTGTAGTCCAGCTA 395
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
240 TCTACTAAAAATACAAAAAGATTAGCTGGGCGTAGTGGGCGGTGGCTGTAGTCCAGCTA 299
QY 396 CTGGGAGGCTGAGCGGGAAGATCGCTTTGAACCCAGTAGGAGAGGTTGAGCTGAGCGG 455
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
300 CTCGAGAGGCTGAGCGAGAGAGATGGCGAACCTGGGAGGCGGAGCTTGCAGTGAGCCA 359
QY 456 AGATAAGAGTCACTCCAGCTCCAGCTGGTGCAGAGAGAGACTCCCTCTCAGAAAAATAA 515
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
360 AGAT-CGCCCACTGCACCTCCAGCTGGGTGCAGAGTGAGACTCCATCTCAAAAAACAA 418
QY 516 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 568
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
419 AACAAAAACAAACAAACAAACAGCTTAAGCTTTCAACTAAGCATCAGGATGGG 471

```

Search completed: November 24, 2002, 09:16:50
Job time : 2194 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2002, 06:08:16 ; Search time 292 Seconds
(without alignments)
7712.323 Million cell updates/sec

Title: US-09-784-423-32

Perfect score: 1000

Sequence: 1 GGCTGACCTTATCTCTCT.....CAGATGATGACCGCGCTGC 1000

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002:*

1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	999	99.9	1000	AA227732	Human DNA marker C
C 2	320.2	32.0	7017	AA137408	Human musculoskele
C 3	320.2	32.0	20522	AA137409	Human musculoskele
C 4	320.2	32.0	92638	ABQ88096	Human osteoblast d
C 5	307	30.7	17424	ASL68122	Ovary cancer relat
C 6	306.8	30.7	2007	AAK73932	Human immune/haema
C 7	306.8	30.7	2007	AAK73933	Human immune/haema
C 8	306.8	30.7	22680	AAK66308	Human immune/haema
C 9	306.8	30.7	22680	AAK73334	Human immune/haema

C 10	306.8	30.7	22680	22	AAK73344	Human immune/haema
C 11	306.8	30.7	22680	22	AAK73625	Human immune/haema
C 12	306.8	30.7	22680	22	AAK73847	Human immune/haema
C 13	306.8	30.7	22680	22	AAK73934	Human immune/haema
C 14	306.8	30.7	22680	22	AAK78350	Human immune/haema
C 15	306.4	30.6	4243	22	ABAL6762	Human nervous syst
C 16	305.6	30.6	15682	22	AAK89506	Human digestive sy
C 17	305.6	30.6	81800	24	ABK84756	Human cDNA differe
C 18	304.4	30.4	13919	24	ABK86218	DNA encoding AIP-1
C 19	304.4	30.4	13919	24	ABK85220	AIP-1/FLASH promot
C 20	304.4	30.4	13919	24	ABK86221	AIP-1/FLASH promot
C 21	304.4	30.4	21404	24	ABK86229	Ovary cancer relat
C 22	301.4	30.1	17424	24	ABL68122	Kidney cancer relat
C 23	301.2	30.1	49634	24	ABL68647	Human transporter
C 24	300.6	30.1	32816	24	ABK49701	Human immune/haema
C 25	299.6	30.0	52616	22	AAK70459	Human immune/haema
C 26	299.6	30.0	52616	22	AAK78930	Human immune/haema
C 27	298	29.8	12221	22	AAK71585	DNA encoding novel
C 28	297.2	29.7	24533	22	AAK71589	Human immune/haema
C 29	296.6	29.7	37314	22	AAK71358	DNA encoding Aldeh
C 30	296.6	29.7	46765	24	AAK59306	Human osteoblast d
C 31	296.6	29.7	172570	24	ABQ88207	Human reproductive
C 32	296.4	29.6	5281	22	AAK05101	Human testicular a
C 33	296.4	29.6	5281	23	ABL97994	Human reproductive
C 34	296.4	29.6	5284	22	AAK05100	Human testicular a
C 35	296.4	29.6	5284	23	ABL97993	Human cDNA differe
C 36	296	29.6	12263	24	ABK84514	Human proto-oncoge
C 37	296	29.6	15297	24	ABK84501	Human secreted pro
C 38	295	29.5	1166	22	AAK19172	Human immune/haema
C 39	295	29.5	21596	22	AAK71582	Human immune/haema
C 40	294.8	29.5	30393	22	AAK67239	Human immune/haema
C 41	294.4	29.4	19490	22	AAK70177	Human nervous syst
C 42	294.2	29.4	32249	22	ABAL7155	Gene #3429 used to
C 43	293.8	29.4	99014	24	ABN6931	Human immune/haema
C 44	293.6	29.4	48045	22	AAK84730	Human immune/haema
C 45	293.6	29.4	48045	22	AAK85984	Human immune/haema

ALIGNMENTS

RESULT 1
AAZ27732
ID AAZ27732 standard; DNA; 1000 BP.
XX
AC AAZ27732;
XX
DT 23-DEC-1999 (first entry)
XX
DE Human DNA marker clone S132.
XX
KW Tandem repeat sequence; DNA isolation; intermediate tandem repeat;
KW IIR sequence; pentanucleotide tandem repeat; stutter artifact;
KW DNA typing; DNA profiling; linkage analysis; criminal justice;
KW paternity testing; animal lineage analysis; microsatellite loci;
KW polymorphism detection; ds.
XX
OS Homo sapiens.
XX
PN WO9940194-A1.
XX
PD 12-AUG-1999.
XX
PF 04-FEB-1999; 99WO-0502345.
XX
PR 04-FEB-1998; 98US-0018584.
XX
PA (PROM-) PROMEGA CORP.
XX
PI Schumm JW, Bacher JW;
XX
DR WPI; 1999-590696/50.
XX

Isolating DNA containing intermediate tandem repeat sequences, useful in DNA profiling -

Claim 22; Page 78; 11pp; English.

This sequence represents a human DNA marker clone used in the method of the invention. The method is for isolating a fragment of DNA containing an intermediate tandem repeat (ITR) sequence using hybridization selection, and comprises: (a) providing several DNA fragments, at least one of which contains an ITR sequence, a region of the DNA fragment which contains at least one repeat unit consisting of a sequence of five, six or seven bases repeated in tandem at least two times; (b) providing a stationary support having at least one oligonucleotide associated with it, where the oligonucleotide includes a sequence of nucleotides which is complementary to a portion of the ITR sequence; and (c) combining the DNA fragments with the support under conditions where the DNA fragments support. The method is particularly useful to isolate DNA containing pentanucleotide tandem repeat sequences as well as to detect target ITR DNA sequences having a low incidence of stutter artifacts (no more than 2.4%). The method is useful in DNA profiling for linkage analysis, criminal justice, paternity testing and other forensic and medical uses. DNA typing is also useful for confirming the lineage of horses, dogs and other prize animals. The invention overcomes problems related to the use of microsatellite loci in DNA profiling. The method can detect polymorphisms with a low incidence of stutter artifacts, which has previously been a problem in interpreting allelic content of loci. The development of markers based on larger repeat units, enables easier separation of the fragments on electrophoretic gels. This allows the simultaneous analysis of more loci.

Sequence 1000 BP; 302 A; 240 C; 231 G; 226 T; 1 other;

Query Match 99.9%; Score 999; DB 20; Length 1000;
Best Local Similarity 100.0%; Pred. No. 4.7e-196;
Matches 1000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGTGACCTTATCCTCTCTGAACCTCAGTTTCTCATCCGTAAATGAACCTGCTAG 60
Db 1 GGTGTGACCTTATCCTCTCTGAACCTCAGTTTCTCATCCGTAAATGAACCTGCTAG 60
QY 61 ATTGTTGTAATAAATAATGAATAGCTAGCGCGGTGGCTCAGCGCTGAATCCCA 120
Db 61 ATTGTTGTAATAAATAATGAATAGCTAGCGCGGTGGCTCAGCGCTGAATCCCA 120
QY 121 GCATTTAGAGGTGAGAGGGTGGATCATTGAGTTCAGGATTTGAGACCAAGCTG 180
Db 121 GCATTTAGAGGTGAGAGGGTGGATCATTGAGTTCAGGATTTGAGACCAAGCTG 180
QY 181 GCCACACGGTGAACCCCATCTCTACTAAAAATAAAAAATTAGCTNGGGTGGCTGGCT 240
Db 181 GCCACACGGTGAACCCCATCTCTACTAAAAATAAAAAATTAGCTNGGGTGGCTGGCT 240
QY 241 CACACCTGTAAATCCAGCAGCTTTGGGAGGCTGAGACGGTGGATCACTGAAGTCAGGAG 300
Db 241 CACACCTGTAAATCCAGCAGCTTTGGGAGGCTGAGACGGTGGATCACTGAAGTCAGGAG 300
QY 301 TTCAGGCCAGCCTGGGCAACATGGTGAACCCAGCTCTCTACTAAAAATAAAAAATTAG 360
Db 301 TTCAGGCCAGCCTGGGCAACATGGTGAACCCAGCTCTCTACTAAAAATAAAAAATTAG 360
QY 361 CCAGGTGTGGTGGCAGACGGCTGTAGTCCAGCTACTTGGGAGGCTGAGGCGGGAAGATC 420
Db 361 CCAGGTGTGGTGGCAGACGGCTGTAGTCCAGCTACTTGGGAGGCTGAGGCGGGAAGATC 420
QY 421 GCTTGACCCAGTAGGAGAGGTTCAGTGAAGCGGAGATAAGATCACTGCTCCAGCC 480
Db 421 GCTTGACCCAGTAGGAGAGGTTCAGTGAAGCGGAGATAAGATCACTGCTCCAGCC 480
QY 481 TGGGTGACAGCAGCAAGACTCCCTCTCAGAAAAATAAAAAATAAATAAATAAATAA 540
Db 481 TGGGTGACAGCAGCAAGACTCCCTCTCAGAAAAATAAAAAATAAATAAATAAATAA 540

QY 541 AATAAATAAATAAATAAATCTTAAAGGGCTGGCATTTGCTAGCATTATATATGCCAATA 600
Db 541 AATAAATAAATAAATAAATCTTAAAGGGCTGGCATTTGCTAGCATTATATATGCCAATA 600
QY 601 AGTAATAGCTATCAATATATCCACCCCTACCACTGCTGTAATTTAGTTTCTTTTTC 660
Db 601 AGTAATAGCTATCAATATATCCACCCCTACCACTGCTGTAATTTAGTTTCTTTTTC 660
QY 661 ACCCCCCATTAGACTTAAGGCAAGATTTCTACCGCTACTCTCTGTAAATTTCTGGTCT 720
Db 661 ACCCCCCATTAGACTTAAGGCAAGATTTCTACCGCTACTCTCTGTAAATTTCTGGTCT 720
QY 721 GGCCATAGTTGGGTCTCAGTGAACATPGGTGAGTGAATGAGCAATGCAAGGAATCTCC 780
Db 721 GGCCATAGTTGGGTCTCAGTGAACATPGGTGAGTGAATGAGCAATGCAAGGAATCTCC 780
QY 781 AGGCCATCTGGAGCCCTCCAGCGGGTGGTCTGGGTCCCACTCCGACCCCATCTCTGACTC 840
Db 781 AGGCCATCTGGAGCCCTCCAGCGGGTGGTCTGGGTCCCACTCCGACCCCATCTCTGACTC 840
QY 841 GGCCACTGAAAGGTAGAGAGTTCTGGGTCCCACTCCGACCCCATCTCTGACTC 900
Db 841 GGCCACTGAAAGGTAGAGAGTTCTGGGTCCCACTCCGACCCCATCTCTGACTC 900
QY 901 TGCTGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 960
Db 901 TGCTGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 960
QY 961 CAGGACTGCAAGGAGCCCGAGCAAGATGATGACCGCGTGC 1000
Db 961 CAGGACTGCAAGGAGCCCGAGCAAGATGATGACCGCGTGC 1000

RESULT 2
AAL37408/c

ID AAL37408 standard; DNA; 7017 BP.

AC AAL37408;

DT 08-JAN-2002 (first entry)

DE Human musculoskeletal system related polynucleotide SEQ ID NO 3773.

KW Cytostatic; immunosuppressive; nontropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.

OS Homo sapiens.

PN W0200155367-Al.

XX 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US01338.

PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209457.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 11-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 22-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 26-SEP-2000; 2000US-0234998.
PR 27-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249284.
PR 17-NOV-2000; 2000US-0249285.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0250390.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0251989.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251858.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-451937/48.

Isolated polypeptide for treating, preventing and/ or prognosing disorders related to the musculoskeletal system including musculoskeletal cancers and also for testing and detection e.g. diagnosis -

Example 2; SEQ ID NO 3773; 781pp + Sequence Listing; English.

The invention relates to novel genes (AAL34669-AAL37666) and proteins (AB03087-AB04109) associated with the musculoskeletal system useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,

CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC (c) cardiovascular disorders such as myocardial ischaemias; (d) wound
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and
CC parasitic infections.

PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 08-NOV-2000; 2000US-024617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 05-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.

PA
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-451937/48.
 XX

Isolated polypeptide for treating, preventing and/or prognosing
 disorders related to the musculoskeletal system including
 musculoskeletal cancers and also for testing and detection e.g.

PT diagnosis -
 XX
 PS Example 2; SEQ ID NO 3774; 781pp + Sequence Listing; English.
 XX

CC The invention relates to novel genes (AAL34669-AAL37666) and proteins
 CC (AB03087-AB04109) associated with the musculoskeletal system useful
 CC for preventing, treating or ameliorating medical conditions e.g. by
 CC protein or gene therapy. The genes are isolated from a range of human
 CC tissues disclosed in the specification. The nucleic acids, proteins,
 CC antibodies and (ant)agonists are useful in the diagnosis, treatment
 CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
 CC other cancers of the adrenal gland, bone, bone marrow, breast,
 CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune hemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
 CC (c) cardiovascular disorders such as myocardial ischaemia; (d) wound
 CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
 CC and (f) infectious diseases such as viral, bacterial, fungal and
 CC parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 20522 BP; 4797 A; 5197 C; 5103 G; 5425 T; 0 other;

Query Match 32.0%; Score 320.2; DB 22; Length 20522;
 Best Local Similarity 84.6%; Pred. No. 1.3e-56;
 Matches 406; Conservative 0; Mismatches 69; Indels 5; Gaps 4;

QY 85 TAGCTAGGCGGGTGGCTCAGCGCTGTATCCAGACATTTAGAGGTGCAAGAGGGTG 145
 DB 10327 TGGCCAGAGCGGGTGGCTCAGCGCTGTATCCAGACATTTAGAGGTGCAAGAGGGTG 10268
 QY 146 GATCATTGAGGTGAGGAGTTTTCAGACAGCGTGGCCACACAGGTGAAACCCATCTCT 205
 DB 10267 GATCATTGAGGAGAGAG-TTCAGACAGCGTGGCCACACAGGTGAAACCCATCTCT 205
 QY 206 ACTAAATAAATAAATTAGCTGGTGGCTCAGCGCTGTATCCAGACATTTAGAGGTG 10209
 DB 10208 ACTAAATAAATAAATTAGCTGGTGGCTCAGCGCTGTATCCAGACATTTAGAGGTG 10149
 QY 266 GAGGCTGAGAGCGGTGGATCAGCTGAAGTCAGAGGTTCAGGCGCAGCGTGGCAACATGG 325
 DB 10148 GAGGCGCAGCGCGGCGATCA--TGAGTCAAGAGATGGAGCATCTCGCTAACATGG 10091
 QY 326 TGAACCCAGCTCTCTACTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 385
 DB 10090 TGAACCCAGCTCTCTACTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 10031
 QY 386 GTCCAGCTTCTTGGGAGGCTGAGCGGGAAGATTCGCTTGAACCCAGTGGCAGAGGTGG 445
 DB 10030 GTCCAGCTTCTTGGGAGGCTGAGCGGGAAGATTCGCTTGAACCCAGTGGCAGAGGTGG 9971
 QY 446 CAGTGGCGGAGATAAGAGTCACTGCCTCAGCGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 505
 DB 9970 CAGTGGCGGAGATCA-CGACACTGCCTCAGCGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 9912
 QY 506 CAGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 565
 DB 9911 CA-AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 9853

RESULT 4
 ABQ88096
 ID ABQ88096 standard; cDNA; 92638 BP.
 XX
 AC ABQ88096;
 XX
 XX
 XX 18-SEP-2002 (first entry)
 XX Human osteoblast differentiation related cDNA SEQ ID NO 3.
 DE
 XX

KW	Human; osteoblast; stem cell differentiation; bone tissue deposition; osteoporosis; osteopathic; ss.
XX	
OS	Homo sapiens.
XX	
PN	W0200250301-A2.
XX	
PD	27-JUN-2002.
XX	
PF	18-DEC-2001; 2001WC-US48276.
XX	
PR	18-DEC-2000; 2000US-255882P.
PR	24-APR-2001; 2001US-285691P.
XX	
PA	(GENE-) GENE LOGIC INC.
PA	(PROC) PROCTER & GAMBLE CO.
PI	Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A; Mertz L;
PI	
XX	
DR	WPI; 2002-557663/59.
XX	
PT	Use of genes and their expression profiles associated with osteoblast differentiation for screening modulators bone formation, for diagnosing or treating e.g. osteoporosis, or as markers for the differentiation process
PT	
PT	
PT	
XX	
PS	Claim 1; SEQ ID NO 3; 78pp + Sequence Listing; English.
XX	
XX	
CC	The invention relates to genes and their expression profiles are used for:
CC	(a) screening modulators of precursor stem cell differentiation into osteoblasts, or bone tissue deposition;
CC	(b) diagnosing abnormal deposition of bone tissue, abnormal rate of osteoblast formation or osteoporosis; or
CC	(c) treating or monitoring treatment of the conditions cited in (b), or monitoring the progression of bone tissue deposition.
CC	Specific conditions include postmenopausal osteoporosis, glucocorticoid osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-induced abnormalities in bone formation or bone loss, conditions that involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis), skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome or fibrous dysplasia. The present sequence is that of an osteoblast differentiation associated cDNA marker of the invention.
CC	Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 92638 BP; 21460 A; 24572 C; 25235 G; 21371 T; 0 other;
	Query Match 32.0%; Score 320.2; DB 24; Length 92638;
	Best Local Similarity 84.6%; Pred. No. 1.5e-56;
	Matches 406; Conservative 0; Mismatches 69; Indels 5; Gaps 4
QY	86 TAGGCTAGCGGGTGGCTCAGCGCTGTATCCACGACTTTAGAAGTGCAGAGGGTG 145
Db	
69545	TGGCCACAGCGGGTGCTCACACCTGTATCCACGACTTTGAAGCGCGGGTG 69604
QY	146 GATCACTTAGGTCAGGAGTTTGTAGACGAGCTGGCCACACGGTGAACCCCATCTCT 205
Db	
69605	GATCACTTAAGACAAGAG-TTCGAGACGAGCTGGCCACATGGTGAACCTGCTCT 69663
QY	206 ACTAAAATAAAAATTTAGCTNGGTTGGTGGCTGCACACCTGTAAATCCACGACTTTGG 265
Db	
69664	ACTAAAATAAAAAATTGACTGGCGAGTGGCTCAACCTGTATCCACGACTTTGG 69723
QY	266 GAGGCTAGACGGGTGGTCACTTAAGTCAGGAGTTCAAGGCCACCTGGGCAACATGG 325
Db	
69724	GAGCCACAGCGGGGAGATCA--TGAGGTCAGAAGATGAGACCATCTCGTCAACATGG 69781
QY	326 TGAACACGACTCTCTACTAAAAATACAAAAATTTAGCCAGGTGGTGGCAGACCGCTGTA 385
Db	
69782	TGAACCGCGCTCTACTAAAAATACAAAAATTTAGCCAGGATGGTGGCAGACCGCTGCA 69841

PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 01-NOV-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
XX Claim 1; SEQ ID 6459; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 847 sequences given in ABL61664
CC to ABL70110, or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.
XX
XX Sequence 174424 BP; 39582 A; 48304 C; 48535 G; 38003 T; 0 other;
Query Match 30.7%; Score 307; DB 24; Length 174424;
Best Local Similarity 83.1%; Pred. No. 8.2e-54;
Matches 409; Conservative 0; Mismatches 76; Indels 7; Gaps 5;
QY 76 TTAATGGAATGGCTAGCGGGTGGCTCAGCGGTGATCCAGCCTTTAGAGGTC 135
Db 132168 TAAATAAAAGAGCTAGCGGCTGCTCAGCTATATCTTAGCTTTGGAGGCC 132109
QY 136 GAAGAGGTGGATCCTTGGCTGAGTGGAGTTTGGAGCCAGCTGGCCACACGGTGAA 195
Db 132108 GAGGAGGTGGATCCTTGGCTGAGTGGAGTTTGGAGCCAGCTGGCCACACGGTGAA 132050
QY 196 CCCATCTCTACTAAAT-AAAAATAGCTNGGGTGGCTGAGTGGCTGAGTGGTAAAT 252
Db 132049 CCCGCTCTTACAAAATACAAAATATAGCGCGGCGGCTGAGTGGCTGAGTGGTAAAT 131990
QY 253 CCCAGCCTTTGGAGGCTGAGAGCGGTGGATCCTGAGTGGAGTGGAGTGGAGTGGAGT 312
Db 131989 CCGACACCTTTGGAGGCGAGCGGGTGGATCCTGAGTGGAGTGGAGTGGAGTGGAGT 131930
QY 313 CTGGGCAACATGGTGAACACACGCTCTCTACTAAAT-ATACAAAATATAGCGTGGTGGT 371
Db 131929 CTGGGCAACATGGTGAACACACGCTCTCTACTAAAT-ATACAAAATATAGCGTGGTGGT 131870
QY 372 GGCACAGCGTGTAGTCCAGCTACTTGGAGGCTGAGCGGGAAGATCGCTTGAACCCA 431
Db 131869 GGCACAGCGTGTAGTCCAGCTACTTGGAGGCTGAGCGGGAAGATCGCTTGAACCCA 131810
QY 432 GTAGGAGAGTGTGAGTGGAGCGGAGTAAAGTCTGCTGCTCCAGCTGGGTGACAGA 491
Db 131809 GGAGGAGAGTGTGAGTGGAGCGGAGTAAAGTCTGCTGCTCCAGCTGGGTGACAGA 131751
QY 492 GCAAGACTCCCTCTCAG-AAAAATAAAATAAAATAAAATAAAATAAAATAAAATAA 550

Db 131750 GTGAGACTTCATCTCAGAAAAACAAACAAACAAACAAACAAACAAACAAACAA 131691
QY 551 AATAAAATCTTA 562
Db 131690 AACAAAAATTA 131679
RESULT 6
AAK73932/C
ID AAK73932 standard; DNA; 2007 BP.
XX AC AAK73932;
XX DT 07-NOV-2001 (first entry)
XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID No:28744.
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX OS cytosolic; gene therapy; vaccine; metastasis; ds.
XX OS Homo sapiens.
XX PN WO200157182-A2.
XX PD 09-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01354.
XX PR 31-JAN-2000; 2000US-0179065.
XX PR 04-FEB-2000; 2000US-0180628.
XX PR 24-FEB-2000; 2000US-0184664.
XX PR 02-MAR-2000; 2000US-0186350.
XX PR 16-MAR-2000; 2000US-0189874.
XX PR 17-MAR-2000; 2000US-0190076.
XX PR 18-APR-2000; 2000US-0198123.
XX PR 19-MAY-2000; 2000US-0205515.
XX PR 07-JUN-2000; 2000US-0209467.
XX PR 28-JUN-2000; 2000US-0214886.
XX PR 30-JUN-2000; 2000US-0215135.
XX PR 07-JUL-2000; 2000US-0216647.
XX PR 07-JUL-2000; 2000US-0216880.
XX PR 11-JUL-2000; 2000US-0217487.
XX PR 11-JUL-2000; 2000US-0217496.
XX PR 11-JUL-2000; 2000US-0218290.
XX PR 26-JUL-2000; 2000US-0220963.
XX PR 26-JUL-2000; 2000US-0220964.
XX PR 14-AUG-2000; 2000US-0224518.
XX PR 14-AUG-2000; 2000US-0224519.
XX PR 14-AUG-2000; 2000US-0225213.
XX PR 14-AUG-2000; 2000US-0225214.
XX PR 14-AUG-2000; 2000US-0225266.
XX PR 14-AUG-2000; 2000US-0225267.
XX PR 14-AUG-2000; 2000US-0225268.
XX PR 14-AUG-2000; 2000US-0225270.
XX PR 14-AUG-2000; 2000US-0225447.
XX PR 14-AUG-2000; 2000US-0225757.
XX PR 14-AUG-2000; 2000US-0225758.
XX PR 14-AUG-2000; 2000US-0225759.
XX PR 18-AUG-2000; 2000US-0226279.
XX PR 22-AUG-2000; 2000US-0226681.
XX PR 22-AUG-2000; 2000US-0226869.
XX PR 23-AUG-2000; 2000US-0227182.
XX PR 23-AUG-2000; 2000US-0227009.
XX PR 30-AUG-2000; 2000US-0228924.
XX PR 01-SEP-2000; 2000US-0229287.
XX PR 01-SEP-2000; 2000US-0229343.
XX PR 01-SEP-2000; 2000US-0229344.
XX PR 01-SEP-2000; 2000US-0229345.
XX PR 05-SEP-2000; 2000US-0229509.
XX PR 05-SEP-2000; 2000US-0229513.
XX PR 06-SEP-2000; 2000US-0230437.
XX PR 06-SEP-2000; 2000US-0230438.

```
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234223.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0244674.
PR 08-NOV-2000; 2000US-0244675.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.

PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 05-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251889.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2000US-0254097.
PR 05-JAN-2001; 2000US-02559678.

(PHMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483426/52.
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX Disclosure; SEQ ID NO 28744; 307lpp + Sequence Listing; English.
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAK62170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of hematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention.
SQ Sequence 2007 BP; 402 A; 607 C; 570 G; 428 T; 0 other;

Query Match 30.7%; Score 306.8; DB 22; Length 2007;
Best Local Similarity 84.3%; Pred. No. 5.7e-54;
Matches 380; Conservative 0; Mismatches 68; Indels 3; Gaps 3;

QY 86 TAGGCTAGGCGCGGTGCTCAGCGCTGTAATCCAGCACTTTAGAGGTGCGAGAGGTG 145
Db 1212 TAGGCTAGGCGCGGTGCTCAGCGCTGTAATCCAGCACTTTAGAGGTGCGAGAGGTG 145
QY 146 GATCACTTGGGTGAGGTGAGTGTGAGACACGCTGCGGTGCTCACACCTGTAATCCAGCACTTTG 265
Db 1152 GATCACTTGGGTGAGGTGAGTGTGAGACACGCTGCGGTGCTCACACCTGTAATCCAGCACTTTG 265
QY 206 ACTAAATAATAAAATAGCTTGGGTGCGGTGCTCACACCTGTAATCCAGCACTTTG 1034
Db 1093 ACTAAATAATAAAATAGCTTGGGTGCGGTGCTCACACCTGTAATCCAGCACTTTG 1034
QY 266 GAGGCTGAGACGGGTGATCATCCTGAGTCAGGAGTTCAGGGCAGCCTGGGCAATG 325
Db 1033 GAGGCTGAGTGGGT-GATCATTTGGGTGAGGAGTTCAGGGCAGCCTGGGCAATG 325
QY 326 TGAACCAACGCTCTCTACTATAAATAACAAAAATTAGCCAGGTGTGTGGTGGCACACGCTGTA 385
```

Db 974 TGAACCCGCTCTACTAAATATAAATATAGCCGGCGTGGTGCACGGCCTGTA 915
QY 386 GTCCAGCTACTTGGAGGCTGAGCGGAAGAAATCGCTTGAACCCAGTAGGCGAGGTTG 445
Db 914 ATCCGGCTACTCAGGAGCTGAGCGAGGAGATCGCTTGAGCTGGGAGCGGAGGTTG 855
QY 446 CAGTGGCCGAGATAAGAGTCACTGCCTCCAGCTGGGTGACAGAGCAAGACTCCCTCT 505
Db 854 CGGTGAGCCGAGTCA-TGCCACTGCCTCCAGCTGGCGACACAGCAAAACTCTACT 796
QY 506 CAGAAATAAATAAATAAATAAATAA 536
Db 795 CAAAAAAGAAAAAAGAGAGAGAGA 765
RESULT 7
AAK73933/c
ID AAK73933 standard; DNA; 2007 BP.
XX AC AAK73933;
XX XX
XX 07-NOV-2001 (first entry)
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:28745.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytotatic; gene therapy; vaccine; metastasis; ds.
XX OS Homo sapiens.
XX XX
XX PN WO200157182-A2.
XX PD 09-AUG-2001.
XX XX
XX PF 17-JAN-2001; 2001MO-US01354.
XX XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 26-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 02-OCT-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0241826.
PR 08-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.

Db	1093	ACTTAAATACAAAATTAGCCAGGTGGCGTCTATGCCTATAATCCACGACTTTGG	103
QY	266	GAGGCTCAGACGGTGGATCACCTGAAGTGCAGGAGTTCAAGGCCAGCCTGGCGACATGG	325
Db	1033	GAGGCTCAGGTGGGT-GATCATTTGAGTCCAGAGTTAGAGGCCAGTCTGGCCACATGA	975
QY	326	TGAACCACTCTCTACTATAAATACAAAATTAGCCAGGTGGTGGCACACGCTGTGA	385
Db	974	TGAACCCCGTCTCTACTATAAATATAAAATTAGCCGGCGTGGTGGCACGCCCTGTGA	915
QY	386	GTCCCGAGCTACTTGGGAGCTGAGCGGGAAGATCGCTTGACCCAGTAGGCAAGGTTG	445
Db	914	ATCCCGGCTACTTCAGGAGGCTGAGCGAGGAGATCGCTTGAGCCTGGGAGCGGAGTTG	855
QY	446	CAGTGAGCCGAGATAAGAGTCACTGCATCTCCAGCCTGGTGGCACAGCAAGCTCCCTCT	505
Db	854	CGGTGAGCGGAGATCA-TGCCACTGCACTCCAGCCTCGCGGACAGCAAAACTCTATCT	796
QY	506	CAGAAAATAAATAAATAAATAAATAAATAA 536	
Db	795	CAAAAAAAGAAAAAAGAGAGAAGA 765	
RESULT 8			
AAK66308/C			
ID	AAK66308 standard; DNA; 22680 BP.		
XC	AAK66308;		
XX			
DT	06-NOV-2001 (first entry)		
XX			
DE	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:21120.		
KW	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;		
KW	cytostatic; gene therapy; vaccine; metastasis; ds.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200157182-A2.		
XX			
PD	09-AUG-2001.		
XX			
PF	17-JAN-2001; 2001WO-US01354.		
XX			
PR	31-JAN-2000; 2000US-0179065.		
PR	04-FEB-2000; 2000US-0180628.		
PR	24-FEB-2000; 2000US-0184664.		
PR	02-MAR-2000; 2000US-0186350.		
PR	16-MAR-2000; 2000US-0189874.		
PR	17-MAR-2000; 2000US-0190076.		
PR	18-APR-2000; 2000US-0198123.		
PR	19-MAY-2000; 2000US-0205515.		
PR	07-JUN-2000; 2000US-0209467.		
PR	28-JUN-2000; 2000US-0214886.		
PR	30-JUN-2000; 2000US-0215135.		
PR	07-JUL-2000; 2000US-0216647.		
PR	07-JUL-2000; 2000US-0216880.		
PR	11-JUL-2000; 2000US-0217487.		
PR	11-JUL-2000; 2000US-0217496.		
PR	14-JUL-2000; 2000US-0218290.		
PR	26-JUL-2000; 2000US-0220963.		
PR	26-JUL-2000; 2000US-0220964.		
PR	14-AUG-2000; 2000US-0224518.		
PR	14-AUG-2000; 2000US-0224519.		
PR	14-AUG-2000; 2000US-0225213.		
PR	14-AUG-2000; 2000US-0225214.		
PR	14-AUG-2000; 2000US-0225266.		
PR	14-AUG-2000; 2000US-0225267.		
PR	14-AUG-2000; 2000US-0225268.		
PR	14-AUG-2000; 2000US-0225270.		
PR	14-AUG-2000; 2000US-0225447.		
PR	14-AUG-2000; 2000US-0225757.		

PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 23-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0228927.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0232081.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 20-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.

PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Barash SC, Ruben SM;
 PI WPI; 2001-483426/52.
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX Disclosure; SEQ ID NO 21120; 3071pp + Sequence Listing; English.
 XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 22680 BP; 4444 A; 5635 C; 7046 G; 4555 T; 0 other;
 Query Match 30.7%; Score 306.8; DB 22; Length 22680;
 Best Local Similarity 84.3%; Pred. No. 7.3e-54;
 Matches 380; Conservative 0; Mismatches 68; Indels 3; Gaps 3;

QY 86 TAGGCTAGCGCGTGTCTACGCTGTATCCAGCTTTAGAGGTCGAAGAGGTG 145

Db 9074 TAGGCTGGCGCGGTGGCTCAGCGCTGTAATCTAGCACITTTGGAGAGTGGAGGGCA 9015
QY 146 GATCACTTTGAGTTCAGGAGTTTGTAGACACAGCGTGGCCAAACAGGTGAAACCCCATCTCT 205
Db 9014 GATCACTTTGAGTTCAGGAGTTCAGACACAGCGTGGCCAAACAGGTGAAACCCCATCTCT 8956
QY 206 ACTAAATAAAAATTTAGCTNGGTCGGTGGCTCACACCTGTATATCCACACATTGG 265
Db 8955 ACTAAATACAAATTTAGCCAGGTCGGTGGCTCATGCTATATCCACACATTGG 8896
QY 266 GAGGCTAGACGGGTGGATCACTGAAAGTCAGGAGTTCAAGGGCCAGCTGGGCAACATGG 325
Db 8895 GAGGCTAGGTTGGGT-GATCATTTGAGTTCAGGAGTTAGAGGCCACTCTGGCCACATGA 8837
QY 326 TGAACACCACTCTCTACTATAAATAACAAATTTAGCCAGGTTGGTGGCACACGCTGTA 385
Db 8836 TGAACCCCTCTCTACTATAAATAAATAAATTTAGCCGGCTGGTGGCACGCTGTA 8777
QY 386 GTCCCACTACTTGGGAGGCTGAGGGGGAAGATCGCTTGAACCCAGTAGGCAGAGTTG 445
Db 8776 ATCCCGGCTACTCAGGAGGCTGAGGAGGAGATCGCTTGAACCCAGTAGGCAGAGTTG 8717
QY 446 CAGTGACCGAGATGAAGTCACTGCCTCCAGCTGGGTGACAGCAAGACTCCCTCT 505
Db 8716 CGGTGAGCGGAGATCA-TGCCACTGCTCCAGCTGGGCGACACAGCAAACTCTATCT 8658
QY 506 CAGAAATAAATAAATAAATAAATAA 536
Db 8657 CAAAAAAGAAAAAAGAGAAAGA 8627

RESULT 9
AAK73334/c
ID AAK73334 standard; DNA; 22680 BP.
XX AAK73334;
XX
XX
DT 06-NOV-2001 (first entry)
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:28146.
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
OS Homo sapiens.
FN WC200157182-A2.
PD 09-AUG-2001.
XX
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 26-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0231966.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0241826.
PR 08-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.

PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 05-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis.
XX
XX
PS Disclosure; SEQ ID NO 28146; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to prevent,
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.

XX
SQ Sequence 22680 BP; 4444 A; 6635 C; 7046 G; 4555 T; 0 other;
Query Match 30.7%; Score 306.8; DB 22; Length 22680;
Best Local Similarity 84.3%; Pred. NO. 7.3e-54;
Matches 380; Conservative 0; Mismatches 68; Indels 3; Gaps 3;
QY 86 TAGCTAGGCGGGTGGCTCACGCTGTATATCCAGCACCTTTAGAGGTGCGAGAGGGTG 145
|||||
DB 9074 TAGCTGGGCGGGTGGCTCACGCTGTATATCTAGCACCTTTGGAGAGTGAGGAGGGCA 9015
QY 146 GATCACTTGAGGTGAGGAGTTTGGACAGCGCTGGCCAAACGCGTGAACCCCATCTCT 205
|||||
DB 9014 GATCACTTGAGGTGAGGAGT-TCGAGACCAGCTGGCCAAACGCGTGAACCCCTGTCT 8956
QY 206 ACTAAATATAAAATTTAGCTNGGCTGGGTGGCTCACACCTGTATATCCAGCACCTTTGG 265
|||||
DB 8955 ACTAAATATAAAATTTAGCCAGGTGGGTGGCTCATGCTTATATCCAGCACCTTTGG 8896
QY 266 GAGGCTGAGACGGGTGGATCACCCTGAAGTCAGAGTTTCAAGCCAGCGCTGGGCAACATGG 325
|||||
DB 8895 GAGGCTGAGGTGGGT-GATCACTTGAGGTGAGGAGTTAGGCCAGTCTGGCCAAACATGA 8837
QY 326 TGAACACACGCTCTCTACTATAAATACAAAATTTAGCCAGGTGGTGGSCACAGCGCTGTA 385
|||||
DB 8836 TGAACACCGCTCTCTACTATAAATATAAAATTTAGCCAGGTGGTGGSCACAGCGCTGTA 8777
QY 386 GTCCAGCTACTTTGGGAGGCTGAGCGGAAGATTCCTTGAACCCAGTATAGGAGGTTG 445
|||||
DB 8776 ATCCGCGCTACTCAGGAGGCTGAGCGAGGAGATCGCTTGAACCTGGGAGGCGGAGTTG 8717
QY 446 CAGTGAGCGGAGATTAAGATCACTGCTCAGCTCCAGCTGGGTGACAGACAGACTCCCTCT 505
|||||
DB 8716 CGGTGAGCGGAGATCA-TGCCACTGCTCCAGCTCCAGCTGGCGACACACAAACTATCT 8658
QY 506 CAGAAATATAAATAAATAAATAAATAAATAA 536
DB 8657 CAAAAAAGAAAAAAGAAAAAAGAAAAAAG 8627
RESULT 10
AAK73344/c
ID AAK73344 standard; DNA; 22680 BP.
XX
XX AAK73344;
AC AAK73344;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/hematopoietic antigen genomic sequence SEQ ID NO:28156.
XX
XX Human; immune; haematopoietic; immune/hematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
XX WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 18-AUG-2000; 2000US-0225759.
 PR 22-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 25-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.

PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2000US-0254097.
 PR 05-JAN-2001; 2000US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-483426/52.

XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -

XX Disclosure; SEQ ID NO 28156; 3071pp + Sequence Listing; English.

XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)

PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241806.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241836.
PR 01-NOV-2000; 2000US-0241836.
PR 08-NOV-2000; 2000US-02464617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX
XX PS Disclosure; SEQ ID NO 28437; 3071pp + Sequence Listing; English.
XX
XX PS AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX CC

CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
XX SQ Sequence 22680 BP; 4555 A; 7046 C; 6635 G; 4444 T; 0 other;

Query Match 30.7%; Score 306.8; DB 22; Length 22680;
Best Local Similarity 84.3%; Pred. No. 7.3e-54;
Matches 380; Conservative 0; Mismatches 68; Indels 3; Gaps 3;
QY 86 TAGGTAGGCGCGGTGCTCAGCCTGTATCCAGCAGCTTTAGAGGTGCGAGAGGGTG 145
DB 13607 TAGGTAGGCGCGGTGCTCAGCCTGTATCCAGCAGCTTTAGAGGTGCGAGAGGGTG 145
QY 146 GATCAGTTGAGGTCAGGAGTTTGTAGACAGCGCTGCCAACACAGCGTGAACCCCTCTCT 205
DB 13667 GATCAGTTGAGGTCAGGAGT-TCGAGACAGCGCTGCCAACACAGCGTGAACCCCTCTCT 13725
QY 206 ACTAAATAAAAAATTAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 265
DB 13726 ACTAAATAAAAAATTAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 13785
QY 266 GAGGCTGAGACGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 325
DB 13786 GAGGCTGAGGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 13844
QY 326 TGAACACCGCTCTCTACTAAATAATACAAAATAGCCAGGTGGTGGTGGTGGTGGTGG 385
DB 13845 TGAACACCGCTCTCTACTAAATAATACAAAATAGCCAGGTGGTGGTGGTGGTGG 13904
QY 386 CTCCTGAGCTACTTGGGAGGCTGAGGCGGAGGAGTTCAGGCGGAGGCTGGGCGACATGG 445
DB 13905 ATCCCGGCTACTCAGGAGGCTGAGGCGGAGGAGTTCAGGCGGAGGCTGGGCGAGG 13964
QY 446 CAGTGGCCGAGATAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 505
DB 13965 CGGTGAGCGGAGATCA-TGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 14023
QY 506 CAGAAATAAAAAATAAATAAATAAATAA 536
DB 14024 CAGAAATAAAAAATAAATAAATAAATAA 14054

RESULT 12
AAK73847/c
ID AAK73847 standard; DNA; 22680 BP.
XX
AC AAK73847;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:28659.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX

PD 09-AUG-2001.
XX 17-JAN-2001; 2001WO-US01354.
PF 31-JAN-2000; 2000US-0179065.
XX 24-FEB-2000; 2000US-0180628.
PR 04-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232403.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
(HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX


```
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 02-OCT-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.

PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
    useful for preventing, diagnosing and/or treating cancers and
    metastasis
XX Disclosure; SEQ ID NO 28746; 3071pp + Sequence Listing; English.
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
    amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
    activity, and can be used in gene therapy and vaccine production. (I)
    proteins and polynucleotides may be used in the prevention, diagnosis and
    treatment of diseases associated with inappropriate (I) expression. For
    example, they may be used to treat disorders associated with decreased
    expression by rectifying mutations or deletions in a patient's genome
    that affect the activity of (I) by expressing inactive proteins or to
    supplement the patients own production of (I). Additionally, (I)
    polynucleotides may be used to produce the secreted (I). Inserting
    the nucleic acids into a host cell and culturing the cell to express the
    protein. (I) proteins and polynucleotides may be used to prevent,
    diagnose and treat immune/hematopoietic-related diseases, especially
    cancers and cancer metastases of hematopoietic-derived cells. AAK64703
    to AAK87694 represent human immune/hematopoietic antigen genomic
    sequences from the present invention. AAK54942 to AAK54950 and AAK82169
    represent sequences used in the exemplification of the present invention.
XX SQ Sequence 22680 BP; 4444 A; 6635 C; 7046 G; 4555 T; 0 other;

Query Match      30.7%; Score 306.8; DB 22; Length 22680;
Best Local Similarity 84.3%; Pred. No. 7.3e-54;
Matches 380; Conservative 0; Mismatches 66; Indels 3; Gaps 3;

QY 86 TAGGCTAGGCGCGGTGGCTCAGCGCTGTATCCAGACCTTAGAAGGTCGAGAGGGTG 145
Db 9074 TAGGCTGGCGCGGTGGCTCAGCGCTGTATCCAGACCTTAGAAGGTCGAGAGGGTGA 9015
QY 146 GATCAGCTGAGGTCAGGAGTTTTCAGACAGCGCTGGCCACACGGTGAACCCCATCTCT 205
Db 9014 GATCAGCTGAGGTCAGGAGT-TTCAGACAGCGCTGGCCACACGGTGAACCCCATCTCT 8956
QY 206 ACTAAAAATAAAAAATTAGCTNGGTCGCGTGCCTCACACCTGTATCCAGACCTTTGG 265
Db 8955 ACTAAAAATAAAAAATTAGGCGCGGTGGCTCATGCTATATCCAGACCTTTGG 8896
QY 266 GAGGCTGAGACGGGTGGATCACTCAAGTCAGGAGTTCAAGGCGACCTGGGCAACATGG 325
Db 8895 GAGGCTGAGTGGGT-GATCATTTGAGGTCAGGAGTTAGAGGCCACTTGGCCCAACATGA 8837
QY 326 TGAACCCAGCTCTCTACTAAATAATACAAAATTAGCAGGTTGTGGGCACACGCCCTGTA 385
Db 8836 TGAACCCAGCTCTCTACTAAATAATACAAAATTAGCAGGTTGTGGGCACACGCCCTGTA 8777
QY 386 GTCCCGACTACTTGGGAGGTCGAGCGGGAAGAAATCGCTTGAACCCAGTAGGAGAGTTG 445
Db 8776 ATCCCGGCTACTCAGGAGGTCGAGCGGGAAGAAATCGCTTGAACCCAGTAGGAGAGTTG 8717
QY 446 CAGTGAGCCGAGATAAGAGTCACTGCACCTCCAGCTGGGTGACAGAGCAAGACTCCCTCT 505
Db 8716 CGGTGAGCCGAGATCA-TGCCACTGCCTCCAGCTCGGCGGACACAGCAAACTCTATCT 8658
QY 506 CAGAAATAATAATAATAATAATAATAATAA 536
Db 8657 CAAAAAAGAAAAAAGAAAAAAGAGAGAAAGA 8627

RESULT 14
```

AAK78350/c
 ID AAK78350 standard; DNA; 24680 BP.
 XX AC AAK78350;
 XX DT 07-NOV-2001 (first entry)
 XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:33162.
 XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis; ds.
 XX OS Homo sapiens.
 XX PN WO200157182-A2.
 XX PD 09-AUG-2001.
 XX PF 17-JAN-2001; 2001WO-US01354.
 XX PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209457.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225477.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226686.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246533.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.

PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI; 2001-483426/52.
 DR
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Disclosure; SEQ ID NO 33162; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK62170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent, the
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 22680 BP; 4444 A; 6635 C; 7046 G; 4555 T; 0 other;
 Query Match 30.7%; Score 306.8; DB 22; Length 22680;
 Best Local Similarity 84.3%; Pred. No. 7.3e-54;
 Matches 380; Conservative 0; Mismatches 68; Indels 3; Gaps 3;
 QY 86 TAGGCTAGCGCGGTGGCTCAGCGCTGTAATCCAGCAGCTTTAGAGGTGAGAGGGTG 145
 Db 9074 TAGGCTAGCGCGGTGGCTCAGCGCTGTAATCCAGCAGCTTTAGAGGTGAGAGGGCA 9015
 QY 146 GATCAGTTAGGTCAGAGTTTGTAGACCGAGCTGGCCACACCGTGAACCCCATCTCT 205
 Db 9014 GATCAGTTAGGTCAGAGT-TCGAGACCGAGCTGGCCACACCGTGAACCCCATCTCT 8956
 QY 206 ACTAAAAATAAAATAGCTNGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 265
 Db 8955 ACTAAAAATAAAATAGCTNGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 8956
 QY 266 GAGGTCAGAGCGGTGGATCAGCTGAAGTCAGAGTTCAAGCCAGCTGGGCAACMG 325
 Db 8895 GAGGTCAGAGCGGTGGATCAGCTGAAGTCAGAGTTCAAGCCAGCTGGGCAACMG 8837
 QY 326 TGAACACAGCTCTACTATAAATACAAAAATAGCCAGGTGGTGGCCACCGCTGTA 385
 Db 8836 TGAACACAGCTCTACTATAAATACAAAAATAGCCAGGTGGTGGCCACCGCTGTA 8777
 QY 386 GTCCAGCTACTTGGAGGCTGAGCGGAGAGATCGCTTGACCCAGTGGAGAGGTG 445
 Db 8776 ATCCGCGCTACTCAGAGGCTGAGCGGAGAGATCGCTTGAGCTGGGAGGAGGTG 8717
 QY 446 CAGTGAAGCGGATAGAGTCACTGACCTGAGCGTGGTGGTGGTGGTGGTGGTGGTGGT 505

Db 8716 CGGTGAGCGGAGATCA-TGCCACTGCAGCTCCAGCTCGGCGACACAGCAAACTCTATCT 8658
 QY 506 CAGAAATAAAATAAAATAAAATAAAATAAA 536
 Db 8657 CAAAAAAGAAAAAAGAAAAAAGAGAGAGA 8627
 RESULT 15
 ABA16762/c
 ID ABA16762 standard; DNA; 4243 BP.
 XX
 AC ABA16762;
 XX
 XX 23-JAN-2002 (first entry)
 XX Human nervous system related polynucleotide SEQ ID NO 9093.
 DE Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
 XX immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
 KW antiparkinsonian; antischlicking; antianaemic; antiarthritic; cancer;
 KW antirheumatic; hepatotropic; cerebroprotective; antinflammatory;
 KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
 KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
 XX Homo sapiens.
 OS
 XX WC200159063-A2.
 PN
 XX 16-AUG-2001.
 PD
 XX 17-JAN-2001; 2001WO-US01334.
 PF
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 30-JUN-2000; 2000US-0214886.
 PR 07-JUL-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 11-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 23-AUG-2000; 2000US-0227182.
 PR 30-AUG-2000; 2000US-0227009.
 PR 01-SEP-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2002, 07:11:02 ; Search time 3161 Seconds
(without alignments)
9206.835 Million cell updates/sec

Title: US-09-784-423-32
Perfect score: 1000
Sequence: 1 GTGTGACCTATCTCTCT.....CAGATGATGACCGGCTGC 1000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.em.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	999	99.9	1000	6	ARI53982	ARI53982 Sequence
c 2	988	98.8	69964	9	HS496C20	283847 Human DNA s
3	956.4	95.6	138459	2	AC090795	AC090795 Homo sapi
c 4	327.4	32.7	62102	2	AC104363	AC104363 Homo sapi
c 5	326.2	32.6	110000	2	AC092921_2	Continuation (3 of
c 6	325.2	32.5	201197	2	HS424J12	282207 Homo sapien
7	324.8	32.5	97110	9	AC110871	AC110871 Homo sapi
c 8	320.2	32.0	92638	9	AC004596	AC004596 Homo sapi
c 9	319.8	32.0	70693	9	AL591583	AL591583 Human DNA
10	319.2	31.9	106216	9	AL161908	AL161908 Human DNA
11	319	31.9	137489	9	AL161909	AL161909 Human DNA
c 12	319	31.9	152192	2	AC027410	AC027410 Mus muscu
13	319	31.9	159367	2	AC044812	AC044812 Homo sapi
c 14	319	31.9	187990	2	AL590805	AL590805 Homo sapi
15	318.4	31.8	195290	9	AL354795	AL354795 Human DNA
16	317.8	31.8	115756	9	HS462215	AL049795 Human DNA
c 17	317.8	31.8	187414	2	AC087163	AC087163 Homo sapi
c 18	317.8	31.8	191950	2	AC022307	AC022307 Homo sapi
19	317.8	31.8	322840	2	HS47675E8	AL121591 Homo sapi
20	316.8	31.7	175923	2	AC064839	AC064839 Homo sapi
21	316.4	31.6	79528	9	HS466N1	297630 Human DNA s
22	316.2	31.6	215077	9	AC112504	AC112504 Homo sapi
c 23	316	31.6	113372	9	AL161613	AL161613 Human DNA
c 24	315.8	31.6	164429	9	AC087164	AC087164 Homo sapi
c 25	314.6	31.5	157075	9	AC005486	AC005486 Homo sapi
26	314.6	31.5	165021	9	AC107019	AC107019 Homo sapi
27	314.4	31.4	225370	2	AL445222	AL445222 Human DNA
28	314.2	31.4	65489	9	AF108083	AF108083 Homo sapi
29	314.2	31.4	143512	2	AC087632	AC087632 Homo sapi
30	313.8	31.4	158624	2	AC087296	AC087296 Homo sapi
c 31	313.8	31.4	158338	2	AC024720	AC024720 Homo sapi
c 32	313.8	31.4	193537	9	AC016168	AC016168 Homo sapi
c 33	313.8	31.4	213475	9	AC019100	AC019100 Homo sapi
c 34	313.4	31.3	165394	2	AP000877	AP000877 Homo sapi
c 35	313.4	31.3	196511	2	AP000854	AP000854 Homo sapi
c 36	313.2	31.3	153101	2	AC026818	AC026818 Homo sapi
37	313.2	31.3	170908	2	AC022557	AC022557 Homo sapi
38	313.2	31.3	201850	2	AC069476	AC069476 Homo sapi
39	313.2	31.3	251805	2	AL603790	AL603790 Homo sapi
40	313	31.3	113850	9	AL137142	AL137142 Human DNA
c 41	313	31.3	153000	2	AC010934	AC010934 Homo sapi
c 42	312.6	31.3	176845	2	AC023509	AC023509 Homo sapi
c 43	312.6	31.3	186335	2	AC068889	AC068889 Homo sapi
c 44	312.2	31.2	144611	2	AC108860	AC108860 Homo sapi
c 45	312.2	31.2	195623	9	AP004218	AP004218 Homo sapi

ALIGNMENTS

RESULT 1
ARI53982
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

ARI53982
Sequence 32 from patent US 6238863.
ARI53982
GI:15122035
Unknown.
Unknown.
Unclassified.
1 (bases 1 to 1000)
Schumm,J.W. and Bachar,J.W.
Materials and methods for indentifying and analyzing intermediate
tandem repeat DNA markers
Patent: US 6238863-A 32 29-MAY-2001;

linear PAT 08-AUG-2001

FEATURES		Location/Qualifiers			
source	1..1000	/organism="unknown"			
BASE COUNT	302 a	240 c	231 g	226 t	1 others
ORIGIN					
Query Match	99.94%	Score 999	DB 6	Length 1000	
Best Local Similarity	100.0%	Pred. No. 5e-251			
Matches 1000	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
QY	1	GGTGTGACCTTATCTCTCTGAACCTCAGTTCCTCATCCGTAAATGAAAAGCTGCTAG	60		
Db	1	GGTGTGACCTTATCTCTCTGAACCTCAGTTCCTCATCCGTAAATGAAAAGCTGCTAG	60		
QY	61	ATTGTGTTAAAAAAATTAAATGGAATAGCTAGGCGGGTGGCTCAGCCCTGTAATCCCA	120		
Db	61	ATTGTGTTAAAAAAATTAAATGGAATAGCTAGGCGGGTGGCTCAGCCCTGTAATCCCA	120		
QY	121	GCATTTAGAAAGTCGAAGAGGGTGGATCACTTGAGGTTCAGGAGTTCCTGAGACAGCCTG	180		
Db	121	GCATTTAGAAAGTCGAAGAGGGTGGATCACTTGAGGTTCAGGAGTTCCTGAGACAGCCTG	180		
QY	181	GCCAAACAGGTGAACCCCACTCTCTACTAAAAATAAAAAATTAAGTTCAGGACAGCCTG	240		
Db	181	GCCAAACAGGTGAACCCCACTCTCTACTAAAAATAAAAAATTAAGTTCAGGACAGCCTG	240		
QY	241	CACACCTGTAATCCCAGCACCTTTGGGAGCGCTGAGACGGTGGATCACTGAAGTCAGGAG	300		
Db	241	CACACCTGTAATCCCAGCACCTTTGGGAGCGCTGAGACGGTGGATCACTGAAGTCAGGAG	300		
QY	301	TTCAAGGCCAGCCTGGGCAACATGTTGAACCAAGTCTCTACTAAAAATAAAAAATTAAG	360		
Db	301	TTCAAGGCCAGCCTGGGCAACATGTTGAACCAAGTCTCTACTAAAAATAAAAAATTAAG	360		
QY	361	CCAGGTGTGGTGGCACAGCGCTGTAGTCCAGCTACTTGGGAGGCTGAGGCGGAAGTCC	420		
Db	361	CCAGGTGTGGTGGCACAGCGCTGTAGTCCAGCTACTTGGGAGGCTGAGGCGGAAGTCC	420		
QY	421	GCTGAAACCCAGTAGGAGAGTTCAGTGTAGCGCGAGATAAGAGTCAGTGCCTCCAGCC	480		
Db	421	GCTGAAACCCAGTAGGAGAGTTCAGTGTAGCGCGAGATAAGAGTCAGTGCCTCCAGCC	480		
QY	481	TGGGTGCAGAGCAGAGCTCCCTCTCAGAAAATAAAATAAAAAATAAAAAATAAA	540		
Db	481	TGGGTGCAGAGCAGAGCTCCCTCTCAGAAAATAAAATAAAAAATAAAAAATAAA	540		
QY	541	AATAAAAATAAAATAAAATTTCTAAAAAGGCTGGCATTTGCTAGCACTTATATGCCCAAT	600		
Db	541	AATAAAAATAAAATAAAATTTCTAAAAAGGCTGGCATTTGCTAGCACTTATATGCCCAAT	600		
QY	601	AGTAATAGCTATCAATATCCCAACCCCTACCACTGTGCTGAAATTAAGTTCTTTTGTG	660		
Db	601	AGTAATAGCTATCAATATCCCAACCCCTACCACTGTGCTGAAATTAAGTTCTTTTGTG	660		
QY	661	ACCCCCATTAGACTTAAGCAGAAATTCACCGTACTCCTCTGTAATTTCTGTTGTTCT	720		
Db	661	ACCCCCATTAGACTTAAGCAGAAATTCACCGTACTCCTCTGTAATTTCTGTTGTTCT	720		
QY	721	GGCACAATAGTTGGTCTCAGTGAACATGTTGAGTGAATGAGCAAAATGAAGGAATCTCC	780		
Db	721	GGCACAATAGTTGGTCTCAGTGAACATGTTGAGTGAATGAGCAAAATGAAGGAATCTCC	780		
QY	781	AGGCATCTGGGAGCCCTCCAGCGGGTGAGTTGGGAAACTCATAGTCTGTCCTCAAT	840		
Db	781	AGGCATCTGGGAGCCCTCCAGCGGGTGAGTTGGGAAACTCATAGTCTGTCCTCAAT	840		
QY	841	GGCCCACTGAAAGGTAGAGAGTTCTGGGTCCCACCTTCGCAACCCCAATCTCTGACTCAC	900		
Db	841	GGCCCACTGAAAGGTAGAGAGTTCTGGGTCCCACCTTCGCAACCCCAATCTCTGACTCAC	900		
QY	901	TGCTGAAAAATAAATAAATAAATAAATCAACCTTATCGGAGCGCTCCACATGCGTTGC	960		
Db	901	TGCTGAAAAATAAATAAATAAATAAATCAACCTTATCGGAGCGCTCCACATGCGTTGC	960		

```

repeat_region /note="AluJb repeat: matches 5. .312 of consensus"
repeat_region /note="MIR repeat: matches 35. .236 of consensus"
misc_feature /note="L2 repeat: matches 2701. .2748 of consensus"
repeat_region /note="match: GSS: Em:AQ356083"
repeat_region /note="AluSx repeat: matches 1. .312 of consensus"
repeat_region /note="AluSg/x repeat: matches 1. .135 of consensus"
repeat_region /note="MIR repeat: matches 70. .146 of consensus"
misc_feature /note="AluSx repeat: matches 1. .297 of consensus"
repeat_region /note="match: GSS: Em:AQ344529"
repeat_region /note="MIR repeat: matches 80. .191 of consensus"
misc_feature /note="264 copies 2 mer cc 58 conserved"
misc_feature /note="forced join; gap estimated as 1kb by restriction
digest data."
repeat_region /note="forced join. Gap estimated as 1kb by restriction
digest data."
repeat_region /note="25 copies 5 mer ctct 62 conserved"
repeat_region /note="MIR repeat: matches 76. .255 of consensus"
repeat_region /note="MER5A repeat: matches 57. .150 of consensus"
repeat_region /note="MER5A repeat: matches 57. .185 of consensus"
repeat_region /note="L2 repeat: matches 1865. .2127 of consensus"
repeat_region /note="AluSx repeat: matches 28. .305 of consensus"
repeat_region /note="L1MA8 repeat: matches 5393. .5854 of consensus"
repeat_region /note="AluSg repeat: matches 1. .310 of consensus"
repeat_region /note="L1MA8 repeat: matches 5340. .5393 of consensus"
repeat_region /note="AluSx repeat: matches 2. .304 of consensus"
repeat_region /note="AluY repeat: matches 1. .306 of consensus"
repeat_region /note="L1MA8 repeat: matches 5352. .5341 of consensus"
repeat_region /note="L1M4 repeat: matches 3784. .5097 of consensus"
repeat_region /note="L1 repeat: matches 3186. .3382 of consensus"
repeat_region /note="MIR repeat: matches 29. .256 of consensus"
misc_feature /note="MIR2FB repeat: matches 2. .414 of consensus"
repeat_region /note="match: GSS: Em:AQ551353"
repeat_region /note="MER91A repeat: matches 114. .193 of consensus"
repeat_region /note="LTR23 repeat: matches 210. .437 of consensus"
repeat_region /note="AluSg repeat: matches 1. .309 of consensus"
repeat_region /note="AluSg repeat: matches 1. .313 of consensus"
repeat_region /note="BURL repeat: matches 6300. .6367 of consensus"
repeat_region /note="MIR repeat: matches 207. .245 of consensus"

```

```

repeat_region 19081. .19391
repeat_region /note="AluSg repeat: matches 1. .307 of consensus"
repeat_region 19392. .19492
repeat_region /note="MIR repeat: matches 117. .207 of consensus"
repeat_region 19581. .19801
repeat_region /note="MIR repeat: matches 29. .261 of consensus"
repeat_region 20302. .20464
repeat_region /note="AluSg/x repeat: matches 134. .302 of consensus"
repeat_region 20465. .20747
repeat_region /note="AluSg repeat: matches 1. .296 of consensus"
repeat_region 22217. .22508
repeat_region /note="AluSx repeat: matches 14. .305 of consensus"
repeat_region 22764. .22848
repeat_region /note="MIR repeat: matches 164. .249 of consensus"
misc_feature 23209. .23725
misc_feature /note="Cpg island"
repeat_region 23946. .24063
repeat_region /note="MIR repeat: matches 21. .140 of consensus"
repeat_region 25053. .25142
repeat_region /note="MIR repeat: matches 21. .119 of consensus"
repeat_region 25207. .25233
repeat_region /note="9 copies 3 mer gct 96 conserved"
misc_feature 25401. .26038
repeat_region /note="match: GSS: Em:B54652"
repeat_region 25877. .26155
repeat_region /note="AluJb repeat: matches 1. .296 of consensus"
repeat_region 26997. .27304
repeat_region /note="AluJo repeat: matches 1. .311 of consensus"
repeat_region 27326. .27595
repeat_region /note="AluSx repeat: matches 37. .311 of consensus"
repeat_region 28098. .28406
repeat_region /note="AluSg repeat: matches 1. .311 of consensus"
repeat_region 28847. .29128
misc_feature /note="L2 repeat: matches 2362. .2745 of consensus"
repeat_region /note="match: GSS: Em:AQ528941"
repeat_region 30940. .30989
repeat_region /note="25 copies 2 mer gt 100 conserved"
repeat_region 30990. .31054
misc_feature /note="MER41A repeat: matches 400. .464 of consensus"
repeat_region 31120. .31339
repeat_region /note="match: STS: Em:G03731; match: STS: Em:G03731"
repeat_region 34416. .34721
repeat_region /note="AluY repeat: matches 1. .310 of consensus"
repeat_region 34776. .34821
repeat_region /note="L2 repeat: matches 2688. .2733 of consensus"
repeat_region 34944. .35243
repeat_region /note="AluSx repeat: matches 1. .300 of consensus"
repeat_region 35351. .35551
repeat_region /note="MER20 repeat: matches 1. .204 of consensus"
repeat_region 36460. .36544
repeat_region /note="L2 repeat: matches 2665. .2749 of consensus"
repeat_region 37480. .37620
repeat_region /note="MIR repeat: matches 4. .150 of consensus"
repeat_region 37861. .38188
misc_feature /note="AluJo repeat: matches 1. .312 of consensus"
misc_feature /note="match: STS: Em:G28177; match: STS: Em:R54672"
misc_feature /note="complement(38086. .38419)"
repeat_region /note="match: STS: Em:G25989"
repeat_region 38802. .38948

```

```

Query Match 98.8%; Score 988; DB 9; Length 69964;
Best Local Similarity 99.9%; Pred. No. 6.5e-248;
Matches 999; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GGGTGACCTTATCTCTGAACCTCAGTTCTTCATCCGTAATAAATAAAGCTGCTAG 60
|||||
DB 6683 GGGTGACCTTATCTCTGAACCTCAGTTCTTCATCCGTAATAAATAAAGCTGCTAG 6624
|||||
QY 61 ATTGTTGTAATAAATAAATAAATGAATAGCTAGCGGCTGCTACCGCTGTAATCCCA 120
|||||

```

Db	6623	ATGTTGTAATAAATTAATGGAATAGGCTAGCGCGGTGGCTCAGCGCTGAATCCCA	6564
QY	121	GCACCTTTAGAGGCTGAAGAGGGTGGATACACTTGAAGTTCAGGAGTTTTTGACACAGCCTG	180
Db	6563	GCACCTTTAGAGGCTCGAGAGGGTGGATACACTTGAAGTTCAGGAGTTTTTGACACAGCCTG	6504
QY	181	GCACACGGGTGAACCCCATCTCTACTAAAAATAAAAAATAGCTGGGTGGGGTGGCT	240
Db	6503	GCACACAGGTGAACCCCATCTCTACTAAAAATAAAAAATAGCTGGGTGGGGTGGCT	6445
QY	241	CACACCTGTAATCCACGACACTTTGGGAGGCTGAGACGGGTGGATGACCTGAAGTCAGAG	300
Db	6444	CACACCTGTAATCCACGACACTTTGGGAGGCTGAGACGGGTGGATGACCTGAAGTCAGAG	6385
QY	301	TTCAAGGCCAGCGTGGCAACATGCTGAACACAGCTCTCTACTAAAAATCAAAAAATTAG	360
Db	6384	TTCAAGGCCAGCGTGGCAACATGCTGAACACAGCTCTCTACTAAAAATCAAAAAATTAG	6325
QY	361	CGAGGTGGTGGCACACGGCTGTAGTCCAGCTACTTGGAGGCTGAGGCGGAAGAATC	420
Db	6324	CGAGGTGGTGGCACACGGCTGTAGTCCAGCTACTTGGAGGCTGAGGCGGAAGAATC	6265
QY	421	GCTTTGAACCCAGTAGCGAGGGTTCAGTGAAGCGGAGTAAGAGTCACTGCATCCAGCC	480
Db	6264	GCTTTGAACCCAGTAGCGAGGGTTCAGTGAAGCGGAGTAAGAGTCACTGCATCCAGCC	6205
QY	481	TGGGTGACAGACAGACTCCCTCTCAGAAAAATAAATAAATAAATAAATAAATAAATAA	540
Db	6204	TGGGTGACAGACAGACTCCCTCTCAGAAAAATAAATAAATAAATAAATAAATAAATAA	6145
QY	541	AATAAATAAATAAATAATCTTAAAGGGCTGGCATTTGGCTAGCATTATATGCCCCAATA	600
Db	6144	AATAAATAAATAAATAATCTTAAAGGGCTGGCATTTGGCTAGCATTATATGCCCCAATA	6085
QY	601	AGTAATAGCTATCAATATCCCAACCCCTACACACTTGCTGGAATTTAGTTTCTTTTTC	660
Db	6084	AGTAATAGCTATCAATATCCCAACCCCTACACACTTGCTGGAATTTAGTTTCTTTTTC	6025
QY	661	ACCCCCATTAGACTTAAGCAGAACTCTCACCGTACTCCTCTGTAAATTTCTGTTTCT	720
Db	6024	ACCCCCATTAGACTTAAGCAGAACTCTCACCGTACTCCTCTGTAAATTTCTGTTTCT	5965
QY	721	GGCATAAGTTGGTCTCAGTGAACATGGTGAATGAGCAAAATGCAAGGAATCTCC	780
Db	5964	GGCATAAGTTGGTCTCAGTGAACATGGTGAATGAGCAAAATGCAAGGAATCTCC	5905
QY	781	AGGCATCTCGGAGCCCTCCAGCGGGTGAAGTTGGGAACTCATAGTCTGTCTCAAT	840
Db	5904	AGGCATCTCGGAGCCCTCCAGCGGGTGAAGTTGGGAACTCATAGTCTGTCTCAAT	5845
QY	841	GGCCCACTGAAGGTGAGAGTTCTTGGGTCCCACTCCGACCCCACTCTCTGACTCAC	900
Db	5844	GGCCCACTGAAGGTGAGAGTTCTTGGGTCCCACTCCGACCCCACTCTCTGACTCAC	5785
QY	901	TGCTGAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	960
Db	5784	TGCTGAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	5725
QY	961	CAGGACTGCAAGGAGCCACGAGAGATGATGACGGCGTGC	1000
Db	5724	CAGGACTGCAAGGAGCCACGAGAGATGATGACGGCGTGC	5685

6
 5
 4
 3
 2
 1

RESULT 3	AC090795	AC090795	138459 bp	DNA	linear	HTG 11-JUN-2001
	LOCUS	Homo sapiens chromosome 8 clone RP11-254K5				
	DEFINITION	SEQUENCE, 12 unordered pieces.				WORKING DRAFT
	ACCESSION	AC090795				
	VERSION	AC090795.2				
	KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.				
	SOURCE	Homo sapiens.				
	ORGANISM	Homo sapiens.				

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 138459)

Birren,B., Linton,L., Nusbaum,C. and Lander,E. Homo sapiens chromosome 8, clone RFL1-254K5 Unpublished

2 (bases 1 to 138459)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S., Barina,N., Bastien,V., Boguslavsky,I., Boukhalter,B., Brown,A., Camarata,C., Campiano,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodgson,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Glade,S., Goyette,M., GrizHugh,W., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,D., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A., Larocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Taylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,G., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,K., Ruse,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupbach,R., Seaman,S., Severy,P., Sounez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,J., Theodore,J., Travers,M., Travis,N., Trigglio,J., Vassiliev,H., Viell,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (10-MAR-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jun 11, 2001 this sequence version replaced gi:13270664.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center -----

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www.seq.wi.mit.edu>

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information -----

Center project name: L12790

Center clone name: 234_K_5

----- Summary Statistics -----

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 133567 bases at least Q40

Consensus quality: 135811 bases at least Q30

Consensus quality: 136678 bases at least Q20

Insert size: 139000; agarose-fp

Insert size: 137359; sum-of-contigs

Quality coverage: 6.5 in Q20 bases; agarose-fp

Quality coverage: 6.6 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 1411: contig of 1411 bp in length

* 1412 1511: gap of 100 bp

* 1512 2165: contig of 654 bp in length

* 2166 2265: gap of 100 bp

* 2266 3074: contig of 809 bp in length

* 3075 3174: gap of 100 bp

* 3175 4763: contig of 1589 bp in length

* 4764 4863: gap of 100 bp

* 4864 6161: contig of 1298 bp in length

* 6162 6261: gap of 100 bp

```

* 6262 8601: contig of 2340 bp in length
* 8602 8701: gap of 100 bp
* 8702 13139: contig of 4438 bp in length
* 13140 13239: gap of 100 bp
* 13240 17613: contig of 4374 bp in length
* 17614 17713: gap of 100 bp
* 17714 30894: contig of 13181 bp in length
* 30895 30994: gap of 100 bp
* 30995 40051: contig of 9057 bp in length
* 40052 40151: gap of 100 bp
* 40152 95859: contig of 55708 bp in length
* 95860 95959: gap of 100 bp
* 95960 138459: contig of 42500 bp in length.
FEATURES
    source
        1..138459
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="8"
            /map="8"
            /clone="RP11-254K5"
            /clone_lib="RPC1-11 Human Male BAC"
            1..1411
                /note="assembly_fragment"
                clone_end:SP6
                vector_side:left
                1512..2165
                    /note="assembly_fragment"
                    2266..3074
                        /note="assembly_fragment"
                        3175..4763
                            /note="assembly_fragment"
                            4864..6161
                                /note="assembly_fragment"
                                6262..8601
                                    /note="assembly_fragment"
                                    8702..13139
                                        /note="assembly_fragment"
                                        13240..17613
                                            /note="assembly_fragment"
                                            17714..30894
                                                /note="assembly_fragment"
                                                30995..40051
                                                    /note="assembly_fragment"
                                                    40152..95859
                                                        /note="assembly_fragment"
                                                        95960..138459
                                                            /note="assembly_fragment"
                                                            clone_end:7
                                                            vector_side:right
BASE COUNT 38140 a 31260 c 30041 g 37914 t 1104 others
ORIGIN

Query Match          95.6%; Score 956.4; DB 2; Length 138459;
Best Local Similarity 97.8%; Pred. No. 1.3e-239;
Matches 998; Conservative 0; Mismatches 1; Indels 21; Gaps 2;

QY 1 GGGTGACCTTATCCTCTCTGAACCTCAGTTTCTCATCCGTAATAAAGAAAGCTCTAG 60
    |||
Db 33521 GGGTGACCTTATCCTCTCTGAACCTCAGTTTCTCATCCGTAATAAAGAAAGCTCTAG 33580
    |||

QY 61 ATTGTTGTAATAAATAAATGGAATAGGCTAGCGCGGTGCTCAGCCCTGTAATCCCA 120
    |||
Db 33581 ATTGTTGTAATAAATAAATGGAATAGGCTAGCGCGGTGCTCAGCCCTGTAATCCCA 33640
    |||

QY 121 GCACCTTAGAGGTCGAAGAGGTGGATCATTGAGGTGAGGATTTTGAGACGACCTG 180
    |||
Db 33641 GCACCTTAGAGGTCGAAGAGGTGGATCATTGAGGTGAGGATTTTGAGACGACCTG 33700
    |||

QY 181 GCCACACGGTGAACCCCATCTCTACTATAAATAAATAAATAGCTNGGTGGGTGGCT 240
    |||
Db 33701 GCCACACGGTGAACCCCATCTCTACTATAAATAAATAAATAGCT-GGGTGGGTGGCT 33759
    |||

QY 241 CACACCTGTAATCCGACACTTTGGAGGCTGAGAGGGTGGATCACCCTGAAGTCAGGAG 300
    |||

```

```

Db 33760 CACACCTGTAATCCGACACTTTGGAGGCTGAGACGGGTGGATCACCCTGAAGTCAGGAG 33819
    |||
QY 301 TTCAGGCCAGCTGGGCAACATGGTGAACACAGTCTCTACTATAAATAAATAAATAAATAAATAA 360
    |||
Db 33820 TTCAGGCCAGCTGGGCAACATGGTGAACACAGTCTCTACTATAAATAAATAAATAAATAAATAA 33879
    |||
QY 361 CAGAGTGGTGGGCAACAGCGCTGTAGTCCAGCTACTTTGGAGGCTGAGGCGGAAGAATC 420
    |||
Db 33880 CCAGGTGGTGGGCAACAGCGCTGTAGTCCAGCTACTTTGGAGGCTGAGGCGGAAGAATC 33939
    |||
QY 421 GCTTGAACCCAGTAGGAGGCTGACGTGAGCGCGAGATAAGATCAGTCTGACATCCAGCC 480
    |||
Db 33940 GCTTGAACCCAGTAGGAGGCTGACGTGAGCGCGAGATAAGATCAGTCTGACATCCAGCC 33999
    |||
QY 481 TGGGTGACAGGACGAGACTCCCTCTCAGAAA-----ATAAATAA 520
    |||
Db 34000 TGGGTGACAGGACGAGACTCCCTCTCAGAAAATAAATAAATAAATAAATAAATAAATAA 34059
    |||
QY 521 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 580
    |||
Db 34060 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 34119
    |||
QY 581 TAGCACTTATATGCCCCAATAAGTATCAATATCCGCCGCCCTACCACTGTGCTG 640
    |||
Db 34120 TAGCACTTATATGCCCCAATAAGTATCAATATCCGCCGCCCTACCACTGTGCTG 34179
    |||
QY 641 AATATTAGTCTCTTTTGTACCCGCCATTAGACTTTAAGGCAGAGATTTCTCAGGCTACTCC 700
    |||
Db 34180 AATATTAGTCTCTTTTGTACCCGCCATTAGACTTTAAGGCAGAGATTTCTCAGGCTACTCC 34239
    |||
QY 701 TCTGTAATTTCTGGTCTCTGACACATAGTTGGTCTCAGTGAACAATGTTGAGTGAATG 760
    |||
Db 34240 TCTGTAATTTCTGGTCTCTGACACATAGTTGGTCTCAGTGAACAATGTTGAGTGAATG 34299
    |||
QY 761 AGCAATGCAAGGAATCTCCAGGCGCATCTGGGAGGCCCTCCAGCGGGGTGAGTTCGGGAA 820
    |||
Db 34300 AGCAATGCAAGGAATCTCCAGGCGCATCTGGGAGGCCCTCCAGCGGGGTGAGTTCGGGAA 34359
    |||
QY 821 ACTCATAGTCTGCTCTCAATGGCCCACTGAAAGGTAGAGAGTTCTGGGTCCCACTCCGC 880
    |||
Db 34360 ACTCATAGTCTGCTCTCAATGGCCCACTGAAAGGTAGAGAGTTCTGGGTCCCACTCCGC 34419
    |||
QY 881 ACCCCCATCTCTGACTCTACTGCTGAAAAATAAATAAATAAATAAATAAATAAATAAATAA 940
    |||
Db 34420 ACCCCCATCTCTGACTCTACTGCTGAAAAATAAATAAATAAATAAATAAATAAATAAATAA 34479
    |||
QY 941 GAGCTCCCACTGCTTGGCAGGACTGCAAGGAGCCAGCAGAGATGATGACCGCGCTG 1000
    |||
Db 34480 GAGCTCCCACTGCTTGGCAGGACTGCAAGGAGCCAGCAGAGATGATGACCGCGCTG 34539
    |||

```

```

RESULT 4
AC104363 Homo sapiens chromosome 18 clone RP11-42N11 map 18, LOW-PASS
LOCUS AC104363 62102 bp DNA linear HTG 08-DEC-2001
DEFINITION Homo sapiens chromosome 18 clone RP11-42N11 map 18, LOW-PASS
SEQUENCE SAMPLING.
ACCESSION AC104363
VERSION AC104363.1 GI:17426352
KEYWORDS HTG; HTGS-PHASE0.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
    1 (bases 1 to 62102)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Allen,N.,
TITLE Homo sapiens chromosome 18, clone RP11-42N11
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 62102)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
    Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B.,
    Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
    Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,

```


Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galsagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hages, B., Hearford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, K., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, N., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Submitted (08-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE
JOURNAL
COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIGR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L22094
Center clone name: 42_N_11

* NOTE: This record contains 81 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be generic-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 665: contig of 665 bp in length
* 666 765: gap of 100 bp
* 766 1454: contig of 689 bp in length
* 1455 1554: gap of 100 bp
* 1555 2232: contig of 678 bp in length
* 2233 2332: gap of 100 bp
* 2333 3028: contig of 696 bp in length
* 3029 3128: gap of 100 bp
* 3129 3753: contig of 625 bp in length
* 3754 3853: gap of 100 bp
* 3854 4459: contig of 606 bp in length
* 4460 4559: gap of 100 bp
* 4560 5225: contig of 666 bp in length
* 5226 5325: gap of 100 bp
* 5326 5979: contig of 654 bp in length
* 5980 6079: gap of 100 bp
* 6080 6752: contig of 673 bp in length
* 6753 6852: gap of 100 bp
* 6853 7526: contig of 674 bp in length
* 7527 7626: gap of 100 bp
* 7627 8307: contig of 681 bp in length
* 8308 8407: gap of 100 bp
* 8408 9090: contig of 683 bp in length
* 9091 9190: gap of 100 bp
* 9191 9880: contig of 690 bp in length
* 9881 9980: gap of 100 bp
* 9981 10667: contig of 687 bp in length
* 10668 10767: gap of 100 bp
* 10768 11436: contig of 669 bp in length

* 11437 11536: gap of 100 bp
* 11537 12157: contig of 621 bp in length
* 12158 12257: gap of 100 bp
* 12258 12913: contig of 656 bp in length
* 12914 13013: gap of 100 bp
* 13014 13667: contig of 654 bp in length
* 13668 13767: gap of 100 bp
* 13768 14429: contig of 662 bp in length
* 14430 14529: gap of 100 bp
* 14530 15189: contig of 660 bp in length
* 15190 15289: gap of 100 bp
* 15290 15950: contig of 661 bp in length
* 15951 16050: gap of 100 bp
* 16051 16724: contig of 674 bp in length
* 16725 16824: gap of 100 bp
* 16825 17505: contig of 681 bp in length
* 17506 17605: gap of 100 bp
* 17606 18290: contig of 685 bp in length
* 18291 18390: gap of 100 bp
* 18391 19050: contig of 660 bp in length
* 19051 19150: gap of 100 bp
* 19151 19807: contig of 657 bp in length
* 19808 19907: gap of 100 bp
* 19908 20568: contig of 661 bp in length
* 20569 20668: gap of 100 bp
* 20669 21322: contig of 654 bp in length
* 21323 21422: gap of 100 bp
* 21423 22079: contig of 657 bp in length
* 22080 22179: gap of 100 bp
* 22180 22852: contig of 673 bp in length
* 22853 22952: gap of 100 bp
* 22953 23621: contig of 669 bp in length
* 23622 23721: gap of 100 bp
* 23722 24386: contig of 665 bp in length
* 24387 24486: gap of 100 bp
* 24487 25174: contig of 688 bp in length
* 25175 25274: gap of 100 bp
* 25275 25957: contig of 683 bp in length
* 25958 26057: gap of 100 bp
* 26058 26729: contig of 672 bp in length
* 26730 26829: gap of 100 bp
* 26830 27478: contig of 649 bp in length
* 27479 27578: gap of 100 bp
* 27579 28237: contig of 659 bp in length
* 28238 28337: gap of 100 bp
* 28338 29002: contig of 665 bp in length
* 29003 29102: gap of 100 bp
* 29103 29764: contig of 662 bp in length
* 29765 29864: gap of 100 bp
* 29865 30544: contig of 680 bp in length
* 30545 30644: gap of 100 bp
* 30645 31330: contig of 686 bp in length
* 31331 31430: gap of 100 bp
* 31431 32107: contig of 677 bp in length
* 32108 32207: gap of 100 bp
* 32208 32887: contig of 680 bp in length
* 32888 32987: gap of 100 bp
* 32988 33679: contig of 692 bp in length
* 33680 33779: gap of 100 bp
* 33780 34451: contig of 672 bp in length
* 34452 34551: gap of 100 bp
* 34552 35216: contig of 665 bp in length
* 35217 35316: gap of 100 bp
* 35317 35978: contig of 662 bp in length
* 35979 36078: gap of 100 bp
* 36079 36731: contig of 653 bp in length
* 36732 36831: gap of 100 bp
* 36832 37486: contig of 655 bp in length
* 37487 37586: gap of 100 bp
* 37587 38249: contig of 663 bp in length
* 38250 38349: gap of 100 bp
* 38350 39032: contig of 683 bp in length
* 39033 39132: gap of 100 bp

```

* 39133 39815: contig of 683 bp in length
* 39816 39915: gap of 100 bp
* 39916 40576: contig of 661 bp in length
* 40577 40676: gap of 100 bp
* 40677 41356: contig of 680 bp in length
* 41357 41456: gap of 100 bp
* 41457 42141: contig of 685 bp in length
* 42142 42241: gap of 100 bp
* 42242 42899: contig of 658 bp in length
* 42900 42999: gap of 100 bp
* 43000 43650: contig of 651 bp in length
* 43651 43750: gap of 100 bp
* 43751 44412: contig of 662 bp in length
* 44413 44512: gap of 100 bp
* 44513 45169: contig of 657 bp in length
* 45170 45269: gap of 100 bp
* 45270 45925: contig of 656 bp in length
* 45926 46025: gap of 100 bp
* 46026 46695: contig of 670 bp in length
* 46696 46795: gap of 100 bp
* 46796 47481: contig of 686 bp in length
* 47482 47581: gap of 100 bp
* 47582 48185: contig of 604 bp in length
* 48186 48285: gap of 100 bp
* 48286 48964: contig of 679 bp in length
* 48965 49064: gap of 100 bp
* 49065 49751: contig of 687 bp in length
* 49752 49851: gap of 100 bp
* 49852 50518: contig of 667 bp in length
* 50519 50618: gap of 100 bp
* 50619 51290: contig of 672 bp in length
* 51291 51390: gap of 100 bp
* 51391 52048: contig of 658 bp in length
* 52049 52148: gap of 100 bp
* 52149 52826: contig of 678 bp in length

Query Match
Best Local Similarity 32.7%; Score 327.4; DB 2; Length 62102;
Matches 328; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 672 GACTTAAGCAGAAATTCACCGTACTCTCTGTAAATTTCTGTTCTCGGCACATAGTT 731
DB 55292 GACTTAAGCAGAAATTCACCGTCTCTCTGTAAATTTCTGTTCTCGGCACATAGTT 55351

QY 732 GGGTCTCAGTGAACATGGTGAATGAGCAAAATGCAAGGAATCTCCAGGCATCTGG 791
DB 55352 GGGTCTCAGTGAACATGGTGAATGAGCAAAATGCAAGGAATCTCCAGGCATCTGG 55411

QY 792 GAGCCCTCCAGCGGGGTGAGTTCGGGAACTCATAGTCTGTCTTCAATGGCCCACTGAA 851
DB 55412 GAGCCCTCCAGCGGGGTGAGTTCGGGAACTCATAGTCTGTCTTCAATGGCCCACTGAA 55471

QY 852 AGGTAGAGATTCTGGTCCCACTCCGCAACCCCACTCTCTGACTCACTGCTGAAAAAT 911
DB 55472 AGGTAGAGATTCTGGTCCCACTCCGCAACCCCACTCTCTGACTCACTGCTGAAAAAT 55531

QY 912 AATAAATAAATAAATAACACTTATCCGAGGCTCCACATGCTCCGAGGACTGCAAA 971
DB 55532 AATAAATAAATAAATAACACTTATCCGAGGCTCCACATGCTCCGAGGACTGCAAA 55591

QY 972 GGAGCCAGCAGAAATGATGACCGGGGTGC 1000
DB 55592 GGAGCCAGCAGAAATGATGACCGGGGTGC 55620

```

```

RESULT 5
AC092921_2/c
WPCOMMENT
Sequence split into 5 fragments
Fragment Name      Begin      End
AC092921_0        1      110000
AC092921_1        100001   210000
AC092921_2        200001   310000

```

```

AC092921_3      300001      410000
AC092921_4      400001      460778
Continuation (3 of 5) of AC092921 from base 200001 (AC092921 Homo sapiens chromosome
Query Match      32.6%; Score 326.2; DB 2; Length 110000;
Best Local Similarity 81.5%; Pred. No. 1.3e-74;
Matches 401; Conservative 0; Mismatches 89; Indels 2; Gaps 2;

QY 70 AAAAAATTAATGAATAGGCTAGGCGGGTGGCTACGCGCTGTAATCCCGACACTTAG 129
DB 60535 AAAAGAGATAAAGCAATCTCTCTGGCGTGGTGGCTCAGCGCTGTAATCCCGACACTTTGG 60476

QY 130 AAGGTGCGAGAGGGTGATCAGTTCAGGTGAGGAGTTTGTAGACAGAGCTGGCCACACG 189
DB 60475 GAGCGCGAGGATGTGATCAGTTCAGGTGAGGAG-TTCGAGACAGCGCTGGCCACACG 60417

QY 190 GTGAACACCCATCTCTACTAAAAATAAAAAATTAAGCTNGGTGGTGGCTCACACTGT 249
DB 60416 GTGAACACCTCTCTCTACTAAGANATCAAAAAATTAAGCGGGGGTGGTGGCTCACACTGT 60357

QY 250 AATCCAGCACTTTGGGAGGCTGAGACGGGTGGATCACCTGAAGTCAGGAGTTCAAGGCC 309
DB 60356 AATCCAGCACTTTGGGAGGCTGAGACGGGTGGATCACCTGAGTTCAGGAGTTTCGAGACC 60297

QY 310 AGCCTGGGCAACATGTTGAAACACAGCTCTCTACTAAAAATAAAAAATTAAGCCAGGTGTG 369
DB 60296 AGCCTGGGCAACATGTTGAAACACAGCTCTCTACTAAAAATAAAAAATTAAGCCGGGGTGG 60237

QY 370 GTGGCACACGCTGTAGTCCCGAGTCTTGGGAGGCTGAGGCGGGAAGAAATCGCTTCAACC 429
DB 60236 GTGAAGCATGCTGTAAATCCCGAGTCTTGGGAGGCTGAGGCGGGAAGAAATCGCTTCAACC 60177

QY 430 CAGTAGGAGAGGTTGAGTTCAGTTCAGGCGAGATAAGAGTCACTGCTCCAGCTGGGTGACA 489
DB 60176 TCGGAGGCGAGAGTTTCACTGAGCTGAGATCA-TGTCATTGCTACAGACTGGCGGACA 60118

QY 490 GAGCAAGACTCCCTCTCAGAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATA 549
DB 60117 GAGCAACACTCCCACTCAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATA 60058

QY 550 AATAAATAATTCT 561
DB 60057 TAAGATACGCT 60046

RESULT 6
HS424J12/c
LOCUS
DEFINITION
Homo sapiens chromosome X clone RP3-424J12, *** SEQUENCING IN
PROGRESS ***, 15 unordered pieces.
ACCESSION
Z82207.3 GI:12331276
VERSION
HTG; HTGS_PHASE1; HTGS_CANCELLED.
KEYWORDS
human.
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 201197)
AUTHORS
Direct Submission
TITLE
Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
JOURNAL
requests: clonerequest@sanger.ac.uk
On Jan 22, 2001 this sequence version replaced gi:10045116.
COMMENT
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: dj424J12
----- Summary Statistics
Assembly program: XGAP4; version 4.5

```

```

Sequencing vector: M13; M77815; 74% of reads
Sequencing vector: plasmid; L08752; 25% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Chemistry: Dye-terminator ABI; 96% of reads
Chemistry: Dye-primer-amersham; 2% of reads
Chemistry: Dye-primer Big Dye; 0% of reads
Consensus quality: 194333 bases at least Q40
Consensus quality: 196512 bases at least Q30
Consensus quality: 198047 bases at least Q20
Insert size: 199797; sum-of-contigs
Quality coverage: 6.87x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

* 1 4154: contig of 4154 bp in length
 4155 4254: gap of 100 bp
 4255 18703: contig of 14449 bp in length
 18704 18803: gap of 100 bp
 18804 20941: contig of 2138 bp in length
 20942 21041: gap of 100 bp
 21042 34907: contig of 13866 bp in length
 34908 35012: gap of 105 bp
 35013 37508: contig of 2496 bp in length
 37509 37608: gap of 100 bp
 37609 52739: contig of 15131 bp in length
 52740 52839: gap of 100 bp
 52840 56933: contig of 4094 bp in length
 56934 57033: gap of 100 bp
 57034 88245: contig of 31212 bp in length
 88246 88346: gap of 101 bp
 88347 93386: contig of 5040 bp in length
 93387 93486: gap of 100 bp
 93487 123956: contig of 30470 bp in length
 123957 124056: gap of 100 bp
 124057 147917: contig of 23861 bp in length
 147918 148017: gap of 100 bp
 148018 176236: contig of 28219 bp in length
 176237 176336: gap of 100 bp
 176337 187128: contig of 10792 bp in length
 187129 187228: gap of 100 bp
 187229 190824: contig of 3596 bp in length
 190825 190939: gap of 135 bp
 190960 201197: contig of 10238 bp in length.

FEATURES
source

```

location/Qualifiers
1. 201197
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP3-424J12"
/clone_lib="RPC1-3"
1. 4154
/note="assembly_fragment:02486
fragment_chain:1"
4255..18703
/note="assembly_fragment:04764
fragment_chain:1"
18804..20941
/note="assembly_fragment:03282
fragment_chain:2"
21042..34907
/note="assembly_fragment:03624
fragment_chain:2"
35008..37508
/note="assembly_fragment:00340"
37609..52739
/note="assembly_fragment:00671.0"
52840..56933

```

	misc_feature	/note="assembly_fragment:01287.0 57034..88245	
	misc_feature	/note="assembly_fragment:01294" 88346..93386	
	misc_feature	/note="assembly_fragment:01608" 93487..123956	
	misc_feature	/note="assembly_fragment:03234" 124057..147917	
	misc_feature	/note="assembly_fragment:04074" 148018..176236	
	misc_feature	/note="assembly_fragment:04816" 176337..187128	
	misc_feature	/note="assembly_fragment:05488" 187229..190824	
	misc_feature	/note="assembly_fragment:05572.0" 190925..201197	
		/note="assembly_fragment:03867 clone_end:T7	
BASE COUNT	55207 a	43266 c	43712 g
ORIGIN		57504 t	1508
Query Match 32.5%; Score 325.2; DB 2;			
Best Local Similarity 81.7%; Pred. No. 2.6e-74;			
Matches 411; Conservative 0; Mismatches 89;			
QY	69	AAAAAATTAATGGAATAGGCTAGGCGCGGTGGCTCAGCGCTGTG	
Db	106056	AAAGAAAGATGAGCCCTGGCTGGTGGCTGGCTCAGCGCTAT	
QY	129	GAAGTCAAGAGGGTGGATCACTTGAGGTGAGAGTTTTGAGAG	
Db	105996	GGAGCCAAAGCGGGTGCATCACTTGAGTGGTGCAGAG-TTCGAGAC	
QY	189	GGTGAACCCCATCTCTACTAAAAATAAAATTAACCTGNGGGTGG	
Db	105937	GGTGAACCCGTCTCTACTAAATACAAAT-GAOCTGGCAC	
QY	249	TATCCAGCACCTTTGGGAGGCTGAGACGGGTGGATCACTTGAAG	
Db	105878	TATCCAGCACCTTTGGGAGGCTGAGTGGTGGATCACTTGAAG	
QY	309	CAGCTGGCAACATGGTGAACCACTCTCTCTAATAAATAACAA	
Db	105818	CAGCCTGTTCACATGGTGAACCCGCTCTCTACTAAAAATAACAA	
QY	369	GGTGGCACAGCCTCTAGTGTCCAGCTACTTTGGGAGGCTGAGGGGG	
Db	105758	GGTGGCAGGCTCCTGTAGTGTCCAGCTACTCTGGGAGGCTGAGACTG	
QY	429	CCAGTAGGCAGAGTTGCAGTGCACCGAGATAGAGTCACTTGCAC	
Db	105698	CCAGGAGCGGAGGTTGACGCGCGCGGAGATCA-TGCCACTTGCAC	
QY	489	AGAGCAAGACTCCCTCTCAGAAAAATAAATAAATAAATAAATAA	
Db	105639	AGAGCGAGACTCCGTCTCAAAAAAATAAATAAATAAATAAATAA	
QY	549	AAATATAATTTAAAGGGCTG	571
Db	105579	ACAACAAAGCGGAGGAAGATG	105557

RESULT 7

AC110871	AC110871	Homo sapiens 3 BAC RP11-639B4 (Roswell Park Cancer Institute Human BAC Library) complete sequence.	97110 bp	DNA	linear	PRI 30-MAY-2002
LOCUS						
DEFINITION						
ACCESSION						
VERSION	AC110871					
KEYWORDS	AC110871.6	GI:21263157				
SOURCE	HTG.					
ORGANISM	Homo sapiens.					
	Homo sapiens					

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE AUTHORS

1 (bases 1 to 97110)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbacia,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleaveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabris,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., He,X., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louissegh,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Moore,S., Morgan,J., Moorish,T., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,M., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabort,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,D., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,Y., Villalon,B., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstock,G. and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 97110)
Worley,K.C.

Direct Submission
Submitted (16-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 97110)
Worley,K.C.

Direct Submission
Submitted (09-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

4 (bases 1 to 97110)
Worley,K.C.

Direct Submission
Submitted (30-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 30, 2002 this sequence version replaced gi:20452973.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the

Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550), searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES	QUALSTAT-REPORT	
	Location/Qualifiers	source
repeat_region	1..97110	/organism="Homo sapiens"
	/db_xref="taxon:9606"	/chromosomes="3"
repeat_region	/clone="RP11-639B4"	/complement(139..266)
	/rpt_family="MIR"	1776..1806
repeat_region	/rpt_family="AT-rich"	1903..2215
	/rpt_family="AluSx"	4975..5004
repeat_region	/rpt_family="AT-rich"	5256..5568
	/rpt_family="AluSg"	5569..5597
repeat_region	/rpt_family="AluSg"	5598..5632
	/rpt_family="AluSg"	7245..7737
repeat_region	/rpt_family="L2"	7717..7844
	/rpt_family="MIR"	10274..10308
repeat_region	/rpt_family="AT-rich"	11121..11398
	/complement(11424..11534)	11574..13164
repeat_region	/rpt_family="MLT1J"	14267..14338
	/rpt_family="MER52A"	14476..14531
repeat_region	/rpt_family="polypurine"	14645..14734
	/rpt_family="MIR"	15617..15959
repeat_region	/rpt_family="MER58B"	16501..16526

```

repeat_region /rpt_family="(TAA)n"
17061..17331
repeat_region /rpt_family="AluJo"
17332..17385
repeat_region /rpt_family="(TAAA)n"
17825..18125
repeat_region /rpt_family="Alusg"
complement(18371..18737)
repeat_region /rpt_family="MLTik"
19143..19448
repeat_region /rpt_family="Alusg"
19737..20018
repeat_region /rpt_family="Alusx"
20019..20045
repeat_region /rpt_family="(TAAA)n"
complement(20081..20370)
repeat_region /rpt_family="Alusx"
complement(20464..20512)
repeat_region /rpt_family="MADEL"
21922..22002
repeat_region /rpt_family="AT-rich"
22209..22366
repeat_region /rpt_family="L2"
complement(22563..22810)
repeat_region /rpt_family="MIR"
complement(23372..23437)
repeat_region /rpt_family="MIR"
complement(23382..23457)
repeat_region /rpt_family="L2"
23452..23531
repeat_region /rpt_family="MIR"
23583..23843
repeat_region /rpt_family="AlusP"
24087..24327
repeat_region /rpt_family="LIME"
24584..24588
repeat_region /rpt_family="(CA)n"
26741..26778
repeat_region /rpt_family="(CA)n"
complement(27665..27957)
repeat_region /rpt_family="AluJb"
28690..28739
repeat_region /rpt_family="(TA)n"
28740..28793
repeat_region /rpt_family="(CA)n"
29182..29225

Query Match 32.5% Score 324.8; DB 9; Length 97110;
Best Local Similarity 83.1% Pred. No. 3e-74;
Matches 393; Conservative 0; Mismatches 78; Indels 2; Gaps 2;

QY 70 AAAAAATTAATGAATAGCTAGGCGGGTGGCTACGCGCTGTAATCCAGCATTGAG 129
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62866 AAGAGATTAAGCAATCTCTGGCGGGTGGCTACGCGCTGTAATCCAGCATTGAG 129
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 130 AAGTCGAAGAGGTGGATCATTGAGTCAGGAGTTTGGAGACCAAGCTGGCCACAGG 189
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62926 GAGCCGAGGAGTGGATCACTGAGTCAGGAGTTTGGAGACCAAGCTGGCCACAGG 189
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 190 GTGAACCCCATCTCTACTAAATAAATAAATAATAGCTNGGGTGGCTGGCTACACCTGT 249
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62985 GTGAACCCCATCTCTACTAAATAAATAAATAATAGCTNGGGTGGCTGGCTACACCTGT 249
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 250 AATCCAGCAGCTTTGGAGGCTGAGCGGGTGGATCACTGAAGTCAGGAGTTCAAGGCC 309
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 63045 AATCCAGCAGCTTTGGAGGCTGAGCGGGTGGATCACTGAAGTCAGGAGTTCAAGGCC 309
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 310 AGCCTGGGCAACATGGTGAACACAGCTCTCTACTTAAATAAATAAATAATAGCAGGTG 369
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 63105 AGCCTGGGCAACATGGTGAACACAGCTCTCTACTTAAATAAATAAATAATAGCAGGTG 369
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 370 GTGGCAACGCGCTAGTCCCGAGCTACTTGGAGGCTGAGCGGGAAGAATCGCTTGAACC 429
||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 63165 GTGAAGCATGCGCTGTAAATCCAGCTACTCAGGAGGCTGAGGAGAGATCGCTTGAACC 63224
QY 430 CAGTAGGAGAGGTTTGAGTGAGCCGAGATAGAGTCACTGCATCCAGCTGGGTGACA 489
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 63225 TGGAGGCGAGAGGTTTTCAGTGAGCTGAGATCA-TGTCATTGCATACAGCTGGCGACA 63283
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 490 GAGCAAGACTCCCTCTCAGAAAAATAAATAAATAAATAAATAAATAAATAAATAA 542
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 63284 GAGCAACACTCCACTCAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 63336
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
AC004596 92638 bp DNA linear PRI 02-FEB-1999
LOCUS Homo sapiens chromosome 17, clone hCIT.175_E_5, complete sequence.
AC004596
AC004596 AC004596.1 GI:4210521
VERSION HTG.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 92638)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 17, clone hCIT.175_E_5
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 92638)
AUTHORS Birren,B., Fasnath,K., McKernan,K., Nusbaum,C., Richardson,P.,
Lander,E., Allen,N., Baker,J., Baldwin,J., Barna,N., Beckerly,R.,
Benn,J., Boatman,C., Boutwell,C., Brown,A., Byrne,S., Cantu,C.,
Castle,A., Cerny,J., Cooke,P., Daly,M.J., Depayre,E., Devon,K.,
Devan,K., Donelan,L., Duketie,B., Etemadi,S., Ferreira,P.,
FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S.,
Gensheimer,S., Geraigery,K., Gilmartin,T., Grant,G., Gray,D.,
Hagos,B., Harris,K., Horton,L., Howland,J.C., Hui,L., Jacotot,L.,
Kann,L., Linton,L., Macdonald,P., Marquis,N., McEwan,P., McGurk,A.,
Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Nachman,A., Nahf,R., Naylor,J., O'Connor,T., Pavlin,B.,
Peterson,K., Riley,R., Roberts,D., Rollins,G., Rossello,R., Roy,A.,
Shyam,R., Stange-Thomann,N., Stilwell,J., Stone,C., Strickland,C.,
Subramanian,A., Sydney,K., Tang,L., Torruella-Miller,I.,
Vassiliev,H., Vo,A., Wagner,A., Wang,B., Wheeler,J., Wu,Y.,
Ye,W.J., Zhao,J. and Zody,M.
DIRECT SUBMISSION
TITLE Direct Submission
JOURNAL Submitted (19-APR-1998) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 92638)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
Castle,A., Cerny,J., Collangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArrellano,K., Depayre,E., Devon,K., Dewar,K.,
Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C.,
Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
Karacas,A., Lehoczyk,J., Lieu,C., Locke,K., Macdonald,P.,
Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
Naylor,J., Nilloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
Testaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
DIRECT SUBMISSION
TITLE Direct Submission
JOURNAL Submitted (02-FEB-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Feb 2, 1999 this sequence version replaced gi:4204188.
All repeats were identified using RepeatMasker: Smit, A.F.A. &
Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html.
FEATURES
Location/Qualifiers
1..92638
/organism="Homo sapiens"
/db_xref="taxon:9606"

```


Db 69545 TGGGCCACGCGGTGCTCACACCTGTTATCCAGCATTGGAAGCGCGAGGGG 69604
 QY 146 GATCACTTGAGGTCAGAGTTTGGAGCAGGCTGGCCAAACGCGTGAACCCCACTCTCT 205
 Db 69605 GATCACTGAAGACAAGAG-TTCGAGACCAAGCTGGCCCAACAGTGTGAACCTCTCTCT 69663
 QY 206 ACTAAATAAATAAATACCTNGGCTGGGTGCGCTCACACCTGTATCCACGACATTGG 265
 Db 69664 ACTAAATAAATAAATGACTGGCGCAGTGGCTCACACCTGTATCCACGACATTGG 69723
 QY 266 GAGGCTGAGCGGTGATCACTGAAGTCAAGAGTTCAGGCGCAGCTGGCGACATGG 325
 Db 69724 GAGGCGAAGCGGCGAGATCA-TGAGGTCAAGATGAGACCATCTGGCTACATGG 69781
 QY 326 TGAACACCACTCTCTACTAATAAATAAATAAATAGCCAGTGTGGTGGACAGCGCTGTA 385
 Db 69782 TGAACCCCGCTCTCTACTAATAAATAAATAAATAGCCAGGCTGGTGGACAGCGCTGCA 69841
 QY 386 GTCCCAAGCTACTTGGGAGGCTGAGGCGGAAGATCGCTGAACCCAGCTAGGCGAGGTTG 445
 Db 69842 GTCCCAAGCTACTGAGGCGTGAAGCAGGAGATAGCTGAACCCAGGCGGAGGTTG 69901
 QY 446 CAGTGACCGGAGATAGAGTCACTGCACTCCAGCTGGGTGACAGAGCAAGACTCCCTCT 505
 Db 69902 CAGTGACCGGAGATCA-CGACACTGCACTCCAGCTGGGCAAGAGCGAGATCCCTCT 69960
 QY 506 CAGAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 565
 Db 69961 CA-AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 70019

RESULT 9
 AL591583/c
 LOCUS
 DEFINITION Human DNA sequence from clone Rp13-163c4 on chromosome X, complete sequence.
 ACCESSION AL591583
 VERSION AL591583.14 GI:20338452
 KEYWORDS HFG.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Submitted (25-APR-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 On Apr 29, 2002 this sequence version replaced gi:16973107.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone. Note that the only a small overlap as described above.
 This sequence was finished as described above.
 regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred/quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SwissProt; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/ChrX

Rp13-163c4 is from the library RPI-13.1 constructed by the group of Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pBac3.6.

FEATURES
 source

1. 70693
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="X"
 /clone="Rp13-163c4"
 /clone_lib="RPI-13.1"
 BASE COUNT 18850 a 16494 c 16633 g 18716 t
 ORIGIN

Query Match 32.0%; Score 319.8; DB 9; Length 70693;
 Best Local Similarity 82.4%; Pred. No. 5.8e-73;
 Matches 402; Conservative 0; Mismatches 83; Indels 3; Gaps 3;

QY 69 AAAAAAATAAATAGGCTAGGCGGTGGCTCAGCGCTGTATCCAGCATTGTA 128
 Db 32038 AAAAGAAAGATGAGCCCTGGCTGGTGGCTCAGCGCTATATCCAGCATTG 31979
 QY 129 GAAGTCCAGAGGGTGGATCTAGTGTGAGGTGAGAGTTTGGAGCAGCCTGCCAACAC 188
 Db 31978 GGAGGCCAAGCGGGTCCATCACTGAGGTGAGGAG-TTCGAGACCAAGCTGACCAACAC 31920
 QY 189 GGTGAACCCCACTCTCTACTAAAAATAAATAAATAGCTNGGGTGGTGGCTCAGCTG 248
 Db 31919 GGTGAATCCCGTCTCTACTAAAAATACAAAAAT-GACCTGGCAGCTTGGCTCAGCTG 31861
 QY 249 TAATCCAGCAGCTTTGGGAGGCTGAGAGCGGTGGATCACTCAAGTCAGGAGTTCAAGGC 308
 Db 31860 TAATCCAGCAGCTTTGGGAGGCTGAGGTGGTGGATCACTCAGGTCAAGAGTTCAAGC 31801
 QY 309 CAGCTGGGCAACATGGTGAACACACGCTCTCTACTAAAAATACAAAAATAGCCAGGTG 368
 Db 31800 CAGCTGTTCAACATGGTGAACACCGCTCTCTACTAAAAATACAAAAATAGCCAGCGT 31741
 QY 369 GGTGCACACCGCTGTAGTCCCGCTGTTGGGAGGCTGAGCGGAGAGTGCCTTGAAC 428
 Db 31740 GGTGCAGGCTCTCTGTAGTCCCGCTGTTGGGAGGCTGAGCGGAGAGTGCCTTGAAC 31681
 QY 429 CCAGTAGCAGAGGTTGAGTGGAGCGAGATAGAGTCACTGCACTCCAGCTGGGTGAC 488
 Db 31680 CCAGTAGCAGGTTGAGTGGAGCGAGATCA-TGCCACTGCACTCCAGCTGGGTGAC 31622
 QY 489 AGAGCAGACTCCCTCTCAGAAAAATAAATAAATAAATAAATAAATAAATAAATAA 548
 Db 31621 AGAGCAGACTCCCTCTCAGAAAAATAAATAAATAAATAAATAAATAAATAAATAA 31562
 QY 549 AAAATAAA 556
 Db 31561 CAACAATA 31554

RESULT 10
 AL161908
 LOCUS
 DEFINITION Human DNA sequence from clone Rp11-123K19 on chromosome 9 Contains part of a novel gene, part of a putative novel gene and CpG islands, complete sequence.

ACCESSION AL161908
 VERSION AL161908.13 GI:14148791
 KEYWORDS HFG; CpG island.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 106216)
 Tracey A.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

Submitted (04-DEC-2001) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 On May 18, 2001 this sequence version replaced gi:13872277.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em, EMBL; SW,
 SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP
 database can be found at
<http://www.sanger.ac.uk/Projects/C-elegans/wormpep> This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
 Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr9>
 RP11-123K19 is from the library RP11-11.1 constructed by the group
 of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pBAC3.6
 IMPORTANT: This sequence is not the entire insert of clone
 RP11-123K19 it may be shorter because we sequence overlapping
 sections only once, except for a short overlap.
 The true left end of clone RP11-123K19 is at 1 in this sequence.
 The true left end of clone RP11-489N22 is at 106117 in this
 sequence.

FEATURES

source	Location/Qualifiers
	1..106216
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="9"
	/clone="RP11-123K19"
	/clone_lib="RP11-11.1"
	7..615
	/note="match: GSS: Em:AQ345966"
	104..325
repeat_region	/note="MIR repeat: matches 14..244 of consensus"
	326..636
repeat_region	/note="AluX repeat: matches 1..310 of consensus"
	776..945
repeat_region	/note="MIR repeat: matches 77..262 of consensus"
	1561..1654
repeat_region	/note="MIR repeat: matches 20..115 of consensus"
	1716..1834
	/note="MIR repeat: matches 53..184 of consensus"
	1745..1907
misc_feature	/note="match: GSS: Em:AQ740851"
	1909..2512
misc_feature	/note="match: GSS: Em:AQ740851"
	1958..2042
repeat_region	/note="MIR repeat: matches 117..185 of consensus"
	2201..2480
repeat_region	/note="AluJb repeat: matches 1..293 of consensus"
	2481..2780
repeat_region	/note="AluSp repeat: matches 1..301 of consensus"
	2871..3156
repeat_region	/note="AluX repeat: matches 2441..2710 of consensus"
	3234..3518
repeat_region	/note="L2 repeat: matches 9..294 of consensus"
	3848..4133
repeat_region	/note="AluX repeat: matches 1..295 of consensus"
	4677..5058
repeat_region	/note="MIR repeat: matches 108..547 of consensus"
	5108..5245
repeat_region	/note="AluJc/FRAM repeat: matches 174..311 of consensus"

repeat_region	5319..5629	/note="AluSg repeat: matches 1..310 of consensus"
	5644..5950	
repeat_region	/note="AluX repeat: matches 1..304 of consensus"	
	6172..6490	
misc_feature	/note="match: GSS: Em:AQ820301"	
	6180..6544	
misc_feature	/note="match: GSS: Em:AQ766292"	
	6203..6506	
repeat_region	/note="AluY repeat: matches 1..304 of consensus"	
	6594..6950	
repeat_region	/note="match: GSS: Em:AQ076946"	
	7477..7628	
misc_feature	/note="MER45 repeat: matches 21..173 of consensus"	
	7754..8091	
repeat_region	/note="match: GSS: Em:AQ566465"	
	7837..7930	
repeat_region	/note="MIR repeat: matches 73..191 of consensus"	
	8066..8127	
repeat_region	/note="31 copies 2 mer cc 74% conserved"	
	8303..8549	
repeat_region	/note="L2 repeat: matches 2483..2736 of consensus"	
	8611..8756	
repeat_region	/note="MIR repeat: matches 60..192 of consensus"	
	9046..9089	
repeat_region	/note="11 copies 4 mer aatg 81% conserved"	
	9093..9124	
repeat_region	/note="16 copies 2 mer aa 87% conserved"	
	9149..9282	
repeat_region	/note="AluSq/X repeat: matches 1..133 of consensus"	
	9298..10471	
repeat_region	/note="LMC2 repeat: matches 5133..6328 of consensus"	
	10529..10828	
repeat_region	/note="AluSp repeat: matches 1..301 of consensus"	
	11333..11557	
repeat_region	/note="L1 repeat: matches 4258..4489 of consensus"	
	11689..11779	
repeat_region	/note="MIR repeat: matches 59..151 of consensus"	
	11867..11948	
repeat_region	/note="MIR repeat: matches 174..255 of consensus"	
	11953..12046	
repeat_region	/note="MER81 repeat: matches 20..113 of consensus"	
	12060..12136	
repeat_region	/note="MIR repeat: matches 69..151 of consensus"	
	12687..12861	
repeat_region	/note="MIR repeat: matches 2..185 of consensus"	
	12862..13016	
repeat_region	/note="MIR repeat: matches 67..236 of consensus"	
	13134..13323	
repeat_region	/note="MIR repeat: matches 39..249 of consensus"	
	13729..13954	
repeat_region	/note="L2 repeat: matches 2427..2642 of consensus"	
	15602..15989	
repeat_region	/note="L2 repeat: matches 2299..2736 of consensus"	
	16104..16237	
repeat_region	/note="L2 repeat: matches 1966..2096 of consensus"	
	16699..16855	
repeat_region	/note="LMD2 repeat: matches 5795..5926 of consensus"	
	16856..17167	
repeat_region	/note="AluX repeat: matches 1..310 of consensus"	
	17168..17354	
repeat_region	/note="LMD2 repeat: matches 5926..6121 of consensus"	
	17355..17410	
repeat_region	/note="14 copies 4 mer ctca 76% conserved"	
	17419..17537	
repeat_region	/note="LMD2 repeat: matches 6138..6258 of consensus"	
	17541..18049	
repeat_region	/note="MIR repeat: matches 4..501 of consensus"	
	18052..18487	
repeat_region	/note="HERV repeat: matches 5330..5758 of consensus"	
	18488..18775	
repeat_region	/note="AluSg repeat: matches 2..289 of consensus"	
	18776..22675	


```
/note="HERVL repeat: matches 1356. .5330 of consensus"
22676. .23077
/note="MTIAL repeat: matches 1. .365 of consensus"
23078. .23162
/note="HERVL repeat: matches 1274. .1356 of consensus"
23164. .23475
/note="AlusX repeat: matches 1. .310 of consensus"
23498. .23809
/note="AlusX repeat: matches 1. .311 of consensus"
23810. .23837
/note="14 copies 2 mer at 92% conserved"
23812. .23839
/note="7 copies 4 mer at 92% conserved"
23839. .24114
/note="AluY repeat: matches 24. .291 of consensus"
24127. .24195
/note="Charlies5 repeat: matches 2471. .2551 of consensus"
24196. .24462
/note="AlusQ repeat: matches 1. .263 of consensus"
24468. .25675
/note="HERVL repeat: matches 36. .1276 of consensus"
25676. .25912
/note="AluY repeat: matches 1. .237 of consensus"
25913. .25948
/note="HERVL repeat: matches 2. .36 of consensus"
25974. .26282
/note="AlusX repeat: matches 18. .312 of consensus"
26285. .26544
/note="L1MD2 repeat: matches 6055. .6323 of consensus"
26545. .26623
/note="MADE1 repeat: matches 1. .80 of consensus"
26624. .26638
/note="L1MD2 repeat: matches 6323. .6336 of consensus"
26812. .27045
/note="MER33 repeat: matches 1. .244 of consensus"
27046. .27344
/note="AlusQ repeat: matches 1. .298 of consensus"
27345. .27448
```

Query Match

```
Best Local Similarity 31.9%; Score 319.2; DB 9; Length 106216;
Matches 39%; Conservative 0; Mismatches 94; Indels 2; Gaps 2;

QY 66 TGTAAAAATTAATAGGCTAGCGGCTGCTCAGCCCTGTATCCGACACT 125
DB 29863 TATATATACATATTAAATATCGCAAGTCAGTGGCTCACACTTGTATGCCGACACT 29922

QY 126 TTAGAAGTCGAGAGGGTGGATCATTGAGGTCAGGAGTTTGGAGCAGCCTGGCCAA 185
DB 29923 TTGGGAGACCAAGCAGGTGGATCAGCTGAGATCAGGAG-TTCAAGCTCAGCTGGCCAA 29981

QY 186 CAGCGTGAACCCATCTCTACTAAATAAATAAATAATAGCTNGGTGGGTGGCTCACAC 245
DB 29982 CATGGTGAACCATGTCTTACTTAAATAAATAAATAAATAGCTGGTGGCTGGCTCACGC 30041

QY 246 CTGTATCCAGCACTTTGGGAGGCTGAGACGGGTGATCAGCTCAGTTCAGTTCAGAGTTCAA 305
DB 30042 CTATATCCAGCACTTTGGGAGGCTGAGACGGGTGATGCTTGGAGTTCAGAGTTCAA 30101

QY 306 GGCAGGCTGGGCAACATGTTGAACACCGCTCTCTACTAAATAAATAAATAATAGCAGG 365
DB 30102 GACCAAGCTAGCCATCAATGCGAACCCTCTCTACTAAATAAATAAATAATAGCCAGG 30161

QY 366 TGTGTGGCAGACGCTGTAGTCCAGCTACTTGGAGGCTGAGCGGGAAGATCGCTTG 425
DB 30162 TGTGTGGCATGTGCTGTAGTCCAGCTACTTGGAGGCTGAGCGGGAAGATCAATCACTTG 30221

QY 426 AACCCAGGTCAGAGGTTGCGAGTCAGGATAGAGTCACTGCCTCCAGCTCCGCTGGGT 485
DB 30222 AACCCAGGTCAGAGGTTGCGAGTCAGGATAGAGTCACTGCCTCCAGCTCCGCTGGGT 30280

QY 486 GACAGAGCAGACTCCCTCTCAGAAAAATAAATAAATAAATAAATAAATAAATAAATAA 545
|||||
```

```
Db 30281 GACAGAGCAGACTCCCTCTCAGAACATATAAATAAATAAATAAATAAATAAATAA 30340
QY 546 AATAAATAAATAAAT 558
| | | | | | | |
Db 30341 CAAAAAATAAATT 30353

RESULT 11
LOCUS AL161909 137289 bp DNA linear PRI 21-NOV-2001
DEFINITION Human DNA sequence from clone RP11-146N23 on chromosome 9, complete
sequence.
ACCESSION AL161909
VERSION AL161909.20 GI:17048334
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 137289)
AUTHORS Ramsay, H.
TITLE Direct Submission
JOURNAL Submitted (21-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
hamquerry@sanger.ac.uk
COMMENT On Nov 22, 2001 this sequence version replaced gi:16943970.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: EM, EMBL; SW,
SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9
RP11-146N23 is from the library RPCI-11.1 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-146N23. It may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true left end of clone RP11-146N23 is at 1 in this sequence.
The true left end of clone RP11-513M16 is at 135290 in this
sequence. The true right end of clone RP11-151U10 is at 26774 in
this sequence.
FEATURES
Location/Qualifiers
1. 137289
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone_lib="RP11-146N23"
/clone_lib="RPCI-11.1"
/3293. .53405
/note="Sequence from reads from a short insert library
derived from a single pUC clone. Restriction digest data
confirm the assembly. Sequence from uni-directional dGTP
big dye terminator reads only."
BASE COUNT 37993 a 27666 c 28638 g 42992 t
ORIGIN
```

Query Match 31.9%; Score 319; DB 9; Length 137289;
Best Local Similarity 80.9%; Pred. No. 1e-72;
Matches 420; Conservative 0; Mismatches 95; Indels 4; Gaps 4;

QY 57 CTAGATTGTTTAAAAAATTAAATAGAAATAGCTAGGCGGTGGCTCAGCGCTGTAAT 116
DQ 57 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT
DQ 30365 CTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT
QY 117 CCACACCTTTAGAGGTGCAAGAGGTGGATCATTTGAGTTCAGAGTGTTCGACACAG 176
DQ 117 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT
DQ 30425 CCCACACCTTTGAGGCGGAGCGGCGGATCACTGAGTTCGGAG-TTTGACACACAG 30483
QY 177 CTTGCGCCAAACAGGTGAACCCCATCTCTACTTAAATTAATAAATTTAGTTCGGT 236
DQ 177 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT
DQ 30484 CTGACTTAACATGGAGAACTCCATCTCTACTTAAACACAAAAATTAGCT-GGGTGTGGT 30542
QY 237 GCTCACAACCTGTAATCCACGACTTTGGAGGCTTGAGCGGTGGATCACTGAGTCA 296
DQ 237 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT
DQ 30543 GGCACACACCTGTAACCCAGCAGCTTTGGAGCGCGGAGCGGTGATCATTTGAGTCA 30602
QY 297 SGAGTTCAAGCCAGCTGGGCAACATGTGGAACACGCTCTCTACTTAAATAACAAAA 356
DQ 297 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT
DQ 30603 GGAGTTTGACACACGCTGGCCAACTGGAGAAACCCCTCTCTACTTAAACACAA 30662
QY 357 TTAGCCAGGTGTGGTGGACAGCCTGTAGTCCGAGTACTTGGAGGCTGAGCGGAG 416
DQ 357 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT
DQ 30663 TTAGTGGCGTGGTGGACATCTGTACCCAGCTACTTGGAGGCTGAGCGGAG 30722
QY 417 AATCGTTGAACCCAGTGGGAGGTGCGAGTGGCGGAGATGAAGTCACTGCTCC 476
DQ 417 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT
DQ 30723 AATCGTTGAACCCAGGAGGTTGCGAGTGGAGTGGAGT-TCCGCCACTGCTCC 30781
QY 477 AGCTGGGTGACAGACGAGTCCCTCTCAGA-AAATAAATAAATAAATAAATAA 535
DQ 477 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT
DQ 30782 AGCTGGGACAGAGTGGAGTCTGTCTATAGAAACAAACAAACAAACAAACAA 30841
QY 536 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 574
DQ 536 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT
DQ 30842 AACAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 30880

RESULT 12
AC027410/c
LOCUS
DEFINITION Mus musculus chromosome 11 clone CT7-549E22 map 11, WORKING DRAFT
SEQUENCE, 35 unordered pieces.
AC027410
AC027410.3 GI:8570504
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H.F., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F., Boguslavskiy,L., Boukhalter,B., Brown,A., Burkett,G., Campilano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dekrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Larocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPeeters,R., Melgrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenka,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,

O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severi,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 17, 2000 this sequence version replaced gi:7652038.
All repeats were identified using RepeatMasker.
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

Project Information
Center project name: L7296
Center clone name: 549_E_22

Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 137974 bases at least Q40
Consensus quality: 144343 bases at least Q30
Insert size: 134000; agarose-1p
Insert size: 148792; sum-of-contigs
Quality coverage: 4.3 in Q20 bases; agarose-1p
Quality coverage: 3.9 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 35 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 1254: contig of 1254 bp in length
* 1255 1354: gap of 100 bp
* 1355 2472: contig of 1118 bp in length
* 2473 2572: gap of 100 bp
* 2573 2968: contig of 396 bp in length
* 2969 3068: gap of 100 bp
* 3069 4814: contig of 1746 bp in length
* 4815 4914: gap of 100 bp
* 4915 6933: contig of 2019 bp in length
* 6934 7033: gap of 100 bp
* 7034 8533: contig of 1500 bp in length
* 8534 8633: gap of 100 bp
* 8634 10357: contig of 1724 bp in length
* 10358 10457: gap of 100 bp
* 10458 12457: contig of 2000 bp in length
* 12458 12557: gap of 100 bp
* 12558 14866: contig of 2309 bp in length
* 14867 14966: gap of 100 bp
* 14967 17009: contig of 2043 bp in length
* 17010 17109: gap of 100 bp
* 17110 19394: contig of 2285 bp in length
* 19395 19494: gap of 100 bp
* 19495 21122: contig of 1628 bp in length
* 21123 21222: gap of 100 bp
* 21223 23617: contig of 2395 bp in length
* 23618 23717: gap of 100 bp
* 23718 26167: contig of 2450 bp in length
* 26168 26267: gap of 100 bp
* 26268 28470: contig of 2203 bp in length
* 28471 28570: gap of 100 bp
* 32045: contig of 3475 bp in length


```

VERSION AC044812.3 GI:8076957
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 159367)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
TITLE Mammalia: Eutheria: Primates: Catarrhini; Homiidae; Homo.
JOURNAL Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE 2 (bases 1 to 159367)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 9, clone RP11-146N23
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 159367)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bieda,F.,
Boguslavsky,L., Bouckhalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galaan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Lakocque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (12-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 25, 2000 this sequence version replaced gi:7844577.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Web site: http://www.seq.wi.mit.edu
Center code: WIBR
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: I7174
Center clone name: 146_N.23
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 144681 bases at least Q40
Consensus quality: 152151 bases at least Q30
Consensus quality: 154947 bases at least Q20
Insert size: 165000; agarose-fp
Insert size: 156467; sum-of-contigs
Quality coverage: 4.0 in Q20 bases; agarose-fp
Quality coverage: 4.2 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 30 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1066: contig of 1066 bp in length
* 1067 1166: gap of 100 bp
* 1167 2259: contig of 1093 bp in length
* 2260 2359: gap of 100 bp
* 2360 3711: contig of 1352 bp in length
* 3712 3811: gap of 100 bp
*

```

```

3812 5274: contig of 1463 bp in length
5275 5374: gap of 100 bp
5375 6956: contig of 1582 bp in length
6957 7056: gap of 100 bp
7057 8817: contig of 1761 bp in length
8818 8917: gap of 100 bp
8918 11538: contig of 2621 bp in length
11539 11638: gap of 100 bp
11639 13616: contig of 1978 bp in length
13617 13716: gap of 100 bp
13717 16484: contig of 2768 bp in length
16485 16584: gap of 100 bp
16585 20183: contig of 3599 bp in length
20184 20283: gap of 100 bp
20284 23098: contig of 2815 bp in length
23099 23198: gap of 100 bp
23199 25692: contig of 2494 bp in length
25693 25792: gap of 100 bp
25793 29609: contig of 3817 bp in length
29610 29709: gap of 100 bp
29710 31132: contig of 1423 bp in length
31133 31232: gap of 100 bp
31233 34267: contig of 3035 bp in length
34268 34367: gap of 100 bp
34368 38304: contig of 3937 bp in length
38305 38404: gap of 100 bp
38405 41936: contig of 3532 bp in length
41937 42036: gap of 100 bp
42037 45621: contig of 3585 bp in length
45622 45721: gap of 100 bp
45722 50410: contig of 4689 bp in length
50411 50510: gap of 100 bp
50511 56294: contig of 5784 bp in length
56295 56394: gap of 100 bp
56395 65126: contig of 8732 bp in length
65127 65226: gap of 100 bp
65227 71685: contig of 6459 bp in length
71686 71785: gap of 100 bp
71786 79513: contig of 7728 bp in length
79514 79613: gap of 100 bp
79614 87044: contig of 7431 bp in length
87045 87144: gap of 100 bp
87145 94789: contig of 7645 bp in length
94790 94889: gap of 100 bp
94890 104863: contig of 9974 bp in length
104864 104963: gap of 100 bp
104964 112662: contig of 7699 bp in length
112663 112762: gap of 100 bp
112763 125278: contig of 12516 bp in length
125279 125378: gap of 100 bp
125379 140442: contig of 15064 bp in length
140443 140542: gap of 100 bp
140543 159367: contig of 18825 bp in length.
Location/Qualifiers
1. .159367
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/map="9"
/clone="RP11-146N23"
/clone_lib="RP11-11 Human Male BAC"
1. .1066
/note="assembly_fragment"
1167. .2259
/note="assembly_fragment"
2360. .3711
/note="assembly_fragment"
3812. 5274
/note="assembly_fragment"
5375. .6956
/note="assembly_fragment"
7057. .8817
/note="assembly_fragment"

```

```

FEATURES
Source
1. .159367
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/map="9"
/clone="RP11-146N23"
/clone_lib="RP11-11 Human Male BAC"
1. .1066
/note="assembly_fragment"
1167. .2259
/note="assembly_fragment"
2360. .3711
/note="assembly_fragment"
3812. 5274
/note="assembly_fragment"
5375. .6956
/note="assembly_fragment"
7057. .8817
/note="assembly_fragment"

```

Query Match	31.9%;	Score 319;	DB 2;	Length 159367;
Best Local Similarity	80.9%;	Pred. No. 1e-72;		
Matches 420;	Conservative	0;	Mismatches 95;	Indels 4;
QY 57	CTAGATTGTTGTAA	AAAAATTAAT	TGAATAGCTAGG	CGCGTGGCTCACCGCTGTAAT 116
Db 59074	CTATTTAATTTAATCTTT	TAAATACTCTAGG	CGCGGTGGCAGTGGCTGATCCCTGTAAT 59133	
QY 117	CCAGCAGCTTTAGAG	GTCGGAAGGGTGGT	CACCTTGAGGTCAGAGGATTTTGAGACACG 176	
Db 59134	CCAGCAGCACTTTGGG	AGCGCGGAGCCAGT	CACCTGAGGTCGGAG-TTTGAGACCAG 59192	
QY 177	CTGTGGCCACACAGGT	GAACCCCACTCTCT	ACTATAAATAAATAATAGCTNGGTTGGGT 236	
Db 59193	CTGTGACTACATGAG	AATACTCCATCTCT	ACTATAAACAACAAATAGCT-GGGGTGGGT 59251	
QY 237	GGCTCACACTGTTAAT	CCAGCAGCACTTTGG	AGGCTGTGAGACGGGTGGTCACTCACTGAGTCA 296	
Db 59252	GGCACACACTGTTAAC	CCCAAGCACTTTGGG	AGCGCCGACCGCGGTGGATCAITTTGAGTCA 59311	
QY 297	GGAGTTTCAGGCGCAG	CGCTGGGCAACATG	TGTGAACCAACCGTCTCTACTAAAAATAACAAAA 356	

COMMENT

```

----- Genome Center:
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA550C2
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 186718 bases at least Q40
Consensus quality: 186962 bases at least Q30
Consensus quality: 187136 bases at least Q20
Insert size: 187290; sum-of-contigs
Quality coverage: 9.19x in Q20 bases; sum-of-contigs Quality
coverage: 9.08x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 95038: contig of 95038 bp in length
* 95039 95138: gap of 100 bp
* 95139 100980: contig of 5842 bp in length
* 100981 101080: gap of 100 bp
* 101081 109127: contig of 8047 bp in length
* 109128 109227: gap of 100 bp
* 109228 112743: contig of 3516 bp in length
* 112744 112843: gap of 100 bp

```

* 112844 157875: contig of 45032 bp in length
 * 157876 157975: gap of 100 bp
 * 157976 171203: contig of 13228 bp in length
 * 171204 171303: gap of 100 bp
 * 171304 178816: contig of 7513 bp in length
 * 178817 178916: gap of 100 bp
 * 178917 187990: contig of 9074 bp in length.

FEATURES

source
 1. .187990
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="13"
 /clone="RP11-550C2"
 /clone_lib="RPC1-11.2"
 1. .95038
 /note="assembly_fragment:00652
 fragment_chain:1
 clone_end:T7
 vector_side:left"
 misc_feature
 95139..100980
 /note="assembly_fragment:03708
 fragment_chain:1"
 101081..109127
 /note="assembly_fragment:03469
 fragment_chain:1"
 109228..112743
 /note="assembly_fragment:02579
 fragment_chain:1"
 112844..1157875
 /note="assembly_fragment:00971
 fragment_chain:1"
 157976..171203
 /note="assembly_fragment:01228
 fragment_chain:1"
 171304..178816
 /note="assembly_fragment:03498
 fragment_chain:1"
 178917..187990
 /note="assembly_fragment:02788
 fragment_chain:1
 clone_end:SP6
 vector_side:right"
 misc_feature
 57887 a 39342 c 37957 g 52104 t 700 others

Query Match 31.9%; Score 319; DB 2; Length 187990;
 Best Local Similarity 80.9%; Pred. No. 1.1e-72;
 Matches 420; Conservative 0; Mismatches 95; Indels 4; Gaps 4;
 QY 57 CTAGATTGTTGTAATAAATTAATGATAGGCTAGGCGCGGTGCTCAGCGCTGTAAT 116
 Db 123939 CTATTTAATTATTTCTTATAAACTCTAAGCGCGGTGCTGCTGTAAT 123880
 QY 117 CCAGCAGCTTTAGAGTCGAGAGGGTGGATCAGTTGAGGTCAGAGCTTTTGAGACCA 176
 Db 123879 CCAGCAGCTTTGGAGCGCGAGCGCGGATCACCTGAGTGGGAG-TTTGAGACCA 123821
 QY 177 CTGCGCCACACGGTGAACCCCATCTCTACTAAAAATAAAATTAAGTNGGTCGGGT 236
 Db 123820 CTGCTAACAATGAGAAATCCATCTCTACTAAAAACAAATAGCT-GGGTGGGT 123762
 QY 237 GGCACACCTGTAATCCAGCAGCTTTGGAGGGTGGAGCGGTGATCATTTGAGGTCA 296
 Db 123761 GGCACACCTGTAATCCAGCAGCTTTGGAGGGTGGAGCGGTGATCATTTGAGGTCA 123702
 QY 297 GGAGTTCAAGGCCAGCTGGGCAACATGTTGAACACCACTCTCTACTAAAAATACAAAA 356
 Db 123701 GGAGTTGAGACAGCTGGGCAACATGTTGAACACCACTCTCTACTAAAAATACAAAA 123642
 QY 357 TTAGCAGGTGTGTGGACACAGCGCTGTAGTCCAGCTACTTGGAGGCTGAGGCGGAG 416
 Db 123641 TTAGCTGGCGGTGTGGACACAGCGCTGTAGTCCAGCTACTTGGAGGCTGAGGCGGAG 123582

QY 417 AATCGCTTGAACCCAGTAGGAGAGGTTCAGTACGCGAGATAGAGTCACTGCACTCC 476
 Db 123581 AATCGCTTGAACCCAGTAGGAGAGGTTCAGTACGCGAGATAGAGTCACTGCACTCC 123523
 QY 477 AGCTGGGTGACAGAGCAAGACTCCCTCTCAGA-AAATAAAATAAAATAAAATAAA 535
 Db 123522 AGCTGGGTGACAGAGTGAAGACTCTCTCTATAGAAACAAACAAACAAACAA 123463
 QY 536 AATAAAATAAAATAAAATTTCTAAAGGGTGGCA 574
 Db 123462 AACAAACAAACAAACAAACCTCTCTAATAGTCTTACA 123424

RESULT 15

AL354795 195290 bp DNA linear PRI 08-SEP-2001
 LOCUS Human DNA sequence from clone RP11-174B4 on chromosome 9q21.2-22.1,
 DEFINITION complete sequence.
 ACCESSION AL354795
 VERSION AL354795.17 GI:14272270
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 195290)
 AUTHORS Hammond, S.
 TITLE Direct Submission
 JOURNAL Submitted (08-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk

COMMENT

On May 31, 2001 this sequence version replaced gi:14132885.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em, EMBL; Sw.,
 SWISSPROT; Tr, TREMBL; Wp., WORMPEP; Information on the WORMPEP
 database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
 Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chrg
 RP11-174B4 is from the library RPC1-11.1 constructed by the group
 of Pieter de Jong. For further details see
 http://www.chori.org/daopac/home.htm
 VECTOR: pBAC63.6

This sequence is the entire insert of clone RP11-174B4 The true
 right end of clone RP11-64C9 is at 58287 in this sequence.

FEATURES

source
 1. .195290
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="9"
 /map="q21.2-22.1"
 /clone="RP11-174B4"
 /clone_lib="RPC1-11.1"
 143978..144033
 misc_feature
 /note="Sequence from uni-directional dGTP big dye
 terminator reads only."

BASE COUNT 62521 a 40135 c 37586 g 55048 t
 ORIGIN

```

Query Match      31.8%; Score 318.4; DB 9; Length 195290;
Best Local Similarity 84.8%; Pred. No. 1.5e-72;
Matches 391; Conservative 0; Mismatches 67; Indels 3; Gaps 3;

QY 84 AATAGGCTAGGCGGGTGGCTCAGCGCTGTAAATCCAGCAGCTTTAGAGGTGGAAGAGGG 143
Db 89288 AATGGCGGGTGCAGTGGCTCAGCGCTGTAAATCCAGCAGCTTTGGAGGCGGAGGCGAGG 89347

QY 144 TGGATCAGTTCAGGTCAGGAGTTTGTAGACCCAGCCTGGCCCAACATGATGAACCCCGTCT 203
Db 89348 TGGATCAGTTCAGGTCAGGAG-TTCGAGACCCAGCCTGGCCCAACATGATGAACCCCGTCT 89406

QY 204 CTACTAAAAATA-AAAAATTAGCTNGGTCGGGTGGCTCAGCGCTGTAAATCCAGCAGCTT 262
Db 89407 CTACTAAAAATAAAAAATTAGGCTGGGTGGCTCAGCGCTGTAAATCCAGCAGCTT 89466

QY 263 TGGAGGCTGAGACGGGTGGATCACCTGAAGTCAGGAGTTCAGGGCCAGCCTGGGCAACA 322
Db 89467 TGGAGGCGCGAGGAGGTGGATCACTTGGAGTCAGGAGTTCAGAGCCAGCCTGGGCAACA 89526

QY 323 TGGTGAACCACTCTCTACTAAAAATAAAAAATTAGCCAGGTGTGGTGGCAGCGCT 382
Db 89527 TGGTGAACCCCTCTCTCTAAAAATAAAAAATTAGCTGGGTGTGGTGGCAGCGCT 89586

QY 383 GTAGTCCAGCTACTTGGGAGGCTGAGCGGAGAAATCGCTTGAACCCAGTAGGCGAGG 442
Db 89587 GTAATCCAGCTACTCCGAGGCTAAGGAGGAGAACTGGCTTGAACCTGGGAGGCGAGG 89646

QY 443 TTGCAGTGAGCGAGATAAGAGTCACTGCAGCTCCAGCTGGGTGACAGAGCAAGACTCCC 502
Db 89647 TTGCAGTGAGCGAGATCTCA-CCACTGCAGCTCCAGCTGGGTGACAGAGCAAGACTCTG 89705

QY 503 TCTCAGAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 543
Db 89706 TCTCAGAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 89746

```

Search completed: November 24, 2002, 08:57:14
Job time : 4193 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2002, 10:24:32 ; Search time 2187 seconds
(without alignments)
7405.344 Million cell updates/sec

Title: US-09-784-423-32

Perfect score: 1000

Sequence: 1 GGTGTGACCTTATCCTCTCT.....CAGAATGATGACGGGGTGC 1000

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_nam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	224	22.4	337	17 AQ356083	AQ356083 CITBI-El-
c 2	66	6.6	490	17 AQ378402	AQ378402 RPC11-16
c 3	62	6.2	703	17 AG089736	AG089736 Pan trogl
c 4	57	5.7	253	12 BF799037	BF799037 CM2-CI003
c 5	57	5.7	327	9 AI216151	AI216151 qm36d06.x
c 6	57	5.7	444	17 AQ284746	AQ284746 RPC111-87

7	57	5.7	455	9	AI886860	AI886860 wk29e07.x
8	57	5.7	476	13	BM353489	BM353489 ig55q11.x
c 9	57	5.7	558	9	AA402129	AA402129 zu55a08.f
c 10	57	5.7	872	14	BQ646204	BQ646204 AGENCOURT
c 11	57	5.7	928	14	BQ648991	BQ648991 AGENCOURT
c 12	57	5.7	948	9	AL554221	AL554221 AL554221
c 13	57	5.7	1086	14	BM803038	BM803038 AGENCOURT
c 14	55	5.5	480	10	AV718952	AV718952 CTW-HSP-2
c 15	55	5.5	501	17	AQ020028	AQ020028 CTW-HSP-2
c 16	55	5.5	629	13	BG912932	BG912932 602807332
c 17	54	5.4	479	9	AV158222	AV158222 ADU58222
c 18	54	5.4	727	17	AG168389	AG168389 Pan trogl
c 19	54	5.4	1103	13	BM477806	BM477806 AGENCOURT
c 20	53	5.3	280	17	AZ276522	AZ276522 RPC1-23-1
c 21	53	5.3	320	9	AA552202	AA552202 nj99c04.s
c 22	53	5.3	329	12	BF149427	BF149427 18_19 Hum
c 23	53	5.3	351	17	AQ125843	AQ125843 HS-2166_B
c 24	53	5.3	352	9	AA411448	AA411448 zv30e05.s
c 25	53	5.3	491	9	AA455483	AA455483 zx76906.s
c 26	53	5.3	577	10	AW973181	AW973181 EST385279
c 27	53	5.3	691	17	AG012178	AG012178 Homo sapi
c 28	53	5.3	693	17	AG012177	AG012177 Homo sapi
c 29	53	5.3	702	17	AG012190	AG012190 Homo sapi
c 30	53	5.3	705	17	AG012191	AG012191 Homo sapi
c 31	52	5.2	293	13	BJ393524	BJ393524 BJ393524
c 32	52	5.2	293	13	BJ396501	BJ396501 BJ396501
c 33	52	5.2	303	13	BJ366477	BJ366477 BJ366477
c 34	52	5.2	305	13	BJ394919	BJ394919 BJ394919
c 35	52	5.2	521	17	AQ350115	AQ350115 RPC111-11
c 36	52	5.2	596	17	AQ376489	AQ376489 RPC111-14
c 37	52	5.2	601	9	AI909139	AI909139 IL-BW200-
c 38	52	5.2	617	13	BI861844	BI861844 603388872
c 39	52	5.2	643	17	AZ955691	AZ955691 2M0221F23
c 40	52	5.2	865	17	AQ740087	AQ740087 HS_5503_A
c 41	52	5.2	962	12	BE883679	BE883679 601507363
c 42	51	5.1	306	10	AW795350	AW795350 PMO-UM001
c 43	51	5.1	569	17	AZ342219	AZ342219 1M0075K08
c 44	51	5.1	588	17	AZ118261	AZ118261 RPC1-23-4
c 45	51	5.1	606	17	AQ932513	AQ932513 RPC1-23-2

ALIGNMENTS

RESULT 1
AQ356083/c
LOCUS
DEFINITION
CITBI-El-2534J20.TF CITBI-El Homo sapiens genomic clone 2534J20,
337 bp DNA linear GSS 24-JAN-1999
DNA sequence.

ACCESSION
AQ356083
VERSION
GSS.
KEYWORDS
GSS.
SOURCE
human.
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Ventur,J.C.
TITLE
Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
Map Building
JOURNAL
Unpublished (1997)

COMMENT
Other_GSSs: CITBI-El-2534J20.TF
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeetigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.


```

Seq primer: M13-21
Class: BAC ends.
FEATURES
  source
    Location/Qualifiers
      1..337
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="2534J20"
        /clone_lib="CITBI-E1"
        /sex="male"
        /cell_type="sperm"
        /note="Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
          Caltech Human BAC Library D"
BASE COUNT      78 a   70 c   75 g   114 t
ORIGIN
      1..114
        22.4%; Score 224; DB 17; Length 337;
        Best Local Similarity 99.6%; Pred. No. 2.7e-96;
        Matches 274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 414 AAGATCGCTGAACCCAGTAGGAGAGGTTGAGTGGAGAGAGATAGAGTCACTGCAC 473
      |||||||
Db 276 AAGATCGCTGAACCCAGTAGGAGAGGTTGAGTGGAGAGAGATAGAGTCACTGCAC 217
      |||||||
QY 474 TCCAGCTGGGTGACAGAGCAAGACTCCCTCTCAGAAAAATAAAATAAAATAAAAT 533
      |||||||
Db 216 TCCAGCTGGGTGACAGAGCAAGACTCCCTCTCAGAAAAATAAAATAAAATAAAAT 157
      |||||||
QY 534 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 157
      |||||||
Db 156 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 97
      |||||||
QY 594 CCATAAGTAATAGCTATCAATATCCCAACCCCTACCACTGTGCTGAAATAGTTTCT 653
      |||||||
Db 96 CCCAATAAGTAATAGCTATCAATATCCCAACCCCTACCACTGTGCTGAAATAGTTTCT 37
      |||||||
QY 654 TTTTGTACCCGCCATTAGACTTAAGCAGAAATTC 688
      |||||||
Db 36 TTTTGTACCCGCCATTAGACTTAAGCAGAAATTC 2

RESULT 2
LOCUS      AQ378402
DEFINITION AQ378402.TV RPCI-11 Homo sapiens genomic clone RPCI-11-164C12,
DNA sequence.
ACCESSION  AQ378402
VERSION    AQ378402.1 GI:4349425
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
            Zhao, S., Adams M.D., Nierman, W., Malek, J., de Jong, P. and Venter
            J.C.
            Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
            Unpublished (1997)
            Other_GSSs: RPCI11-164C12.TJ
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbeetigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@jeng.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7

```

```

Class: BAC ends.
FEATURES
  source
    Location/Qualifiers
      1..490
        /organism="Homo sapiens"
        /db_xref="GDB:7562651"
        /db_xref="taxon:9606"
        /clone="RPCI-11-164C12"
        /clone_lib="RPCI-11"
        /sex="Male"
        /cell_type="Lymphocytes"
        /note="Vector: pBACE3 6; Site_1: EcoRI; Site_2: EcoRI;
          RPCI11 Human Male BAC Library"
BASE COUNT      164 a   111 c   112 g   103 t
ORIGIN
      1..103
        6.6%; Score 66; DB 17; Length 490;
        Best Local Similarity 100.0%; Pred. No. 1.2e-20;
        Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 346 AAATACAAAATAGCCAGGTGTGGTGGACACGCGCTGTAGTCCAGTACTTGGAGGC 405
      |||||||
Db 325 AAATACAAAATAGCCAGGTGTGGTGGACACGCGCTGTAGTCCAGTACTTGGAGGC 384
      |||||||
QY 406 TGAGGC 411
      |||||||
Db 385 TGAGGC 390

RESULT 3
LOCUS      AG089736
DEFINITION AG089736
ACCESSION  AG089736.1 GI:16641538
KEYWORDS   GSS.
SOURCE     Pan troglodytes male lymphoblast DNA, clone: PTB-089B08.R.
ORGANISM   Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Euthera; Primates; Catarrhini; Hominidae; Pan.
            Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
            Totoki, Y., Watanabe, H. and Sakaki, Y.
            BAC end sequences of Library PTB
            Unpublished
            2 (bases 1 to 703)
            Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
            Totoki, Y., Watanabe, H. and Sakaki, Y.
            Direct Submission
            Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail: chimbes@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
            Tel: 81-45-503-9111, Fax: 81-45-503-9170)
            Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
            PRIMERS
            Sequencing: M13Rev
            LIBRARY
            Vector      : pKS145
            R.Site 1    : SacI
            R.Site 2    : SacI.
            Location/Qualifiers
              1..703
                /organism="Pan troglodytes"
                /db_xref="taxon:9598"
                /clone="PTB-089B08.R"
                /sex="male"
                /cell_type="lymphoblast"
                /clone_lib="PTB Chimpanzee Male BAC Library"
BASE COUNT      164 a   164 c   165 g   209 t
ORIGIN

```

Query Match 6.2%; Score 62; DB 17; Length 703;
 Best Local Similarity 100.0%; Pred. No. 9.7e-19;
 Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 336 TCTCTACTAAAAATACAAAAATTAGCCAGGTGGTGGCCACACGCCCTGTAGTCCAGCTA 395
 |||||
 Db 494 TCTCTACTAAAAATACAAAAATTAGCCAGGTGGTGGCCACACGCCCTGTAGTCCAGCTA 435

QY 396 CT 397
 ||
 Db 434 CT 433

RESULT 4
 BF799037

LOCUS CM2-CI0031-051000-409-d04 CI0031 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BF799037
 VERSION BF799037.1 GI:12128026
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 253)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM2&t2=CM2-CI0031-
 051000-409-d04&t3=2000-10-05&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 21
 High quality sequence stop: 253.

FEATURES
 source
 Location/Qualifiers
 1..253
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CI0031"
 /dev_stage="Adult"
 /note="Organ: colon.ins; Vector: puc18; Site.1: SmaI;
 Site.2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 61 a 69 c 68 g 54 t 1 others
 ORIGIN

Query Match 5.7%; Score 57; DB 12; Length 253;
 Best Local Similarity 100.0%; Pred. No. 2.6e-16;
 Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 GAGACAGCGTGGCCACACGGTGAACCCCATCTCTACTAAAAATAAAATTAGC 225
 |||||

Db 145 GAGACAGCGTGGCCACACGGTGAACCCCATCTCTACTAAAAATAAAATTAGC 201

RESULT 5
 A1216151/c

LOCUS A1216151
 DEFINITION 327 bp mRNA linear EST 22-OCT-1998
 gm36d06.x1 NCI-CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1883915 3'-
 similar to contains Alu repetitive element;contains element L1
 repetitive element ;, mRNA sequence.

ACCESSION A1216151
 VERSION A1216151.1 GI:3785192
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 327)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/dbcrp/image/image.html
 Seq primer: -40UP from Gibco
 High quality sequence stop: 296.

FEATURES
 source
 Location/Qualifiers
 1..327
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="1883915"
 /clone_lib="NCI-CGAP_Lu5"
 /tissue_type="carcinoid"
 /lab_host="DH10B"
 /note="Organ: lung; Vector: pVT3D-Pac (Pharmacia) with a
 modified polylinker; 1st strand cDNA was prepared from
 neuroendocrine lung carcinoid, and was then primed with a
 Not I - oligo(dT) primer. Double-stranded cDNA was ligated
 to Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pVT3 vector. Library is normalized. Library was
 constructed by Bento Soares and M. Fatima Bonaldo."
 BASE COUNT 62 a 93 c 84 g 88 t
 ORIGIN

Query Match 5.7%; Score 57; DB 9; Length 327;
 Best Local Similarity 100.0%; Pred. No. 2.6e-16;
 Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 GAGACAGCGTGGCCACACGGTGAACCCCATCTCTACTAAAAATAAAATTAGC 225
 |||||

Db 207 GAGACAGCGTGGCCACACGGTGAACCCCATCTCTACTAAAAATAAAATTAGC 151
 |||||

RESULT 6
 AQ284746/c

LOCUS AQ284746
 DEFINITION 444 bp DNA linear GSS 04-MAY-1999
 RPC111-8705-TV RPCI-11 Homo sapiens genomic clone RPCI-11-8705, DNA
 sequence.

ACCESSION AQ284746
 VERSION AQ284746.1 GI:3910986
 KEYWORDS GSS.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
ADAMS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 444)
Adams, M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K.,
Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P., and Venter, J.C.
Use of human BAC End Sequences for Sequence-Ready Map Building
Unpublished (1998)
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
Seq primer: T7
Class: BAC ends.

FEATURES
source

Location/Qualifiers
1..444
/organism="Homo sapiens"
/db_xref="GDB:7533364"
/db_xref="taxon:9606"
/clone="RPCI-11-8705"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"
BASE COUNT 93 a 97 c 90 g 164 t
ORIGIN

Query Match 5.7%; Score 57; DB 17; Length 444;
Best Local Similarity 100.0%; Pred. No. 2.5e-16;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 169 GAGACCGCTGGCCACACGGTGAACCCCTCTCTACTATAAAATTAAGC 225
Db 60 GAGACCGCTGGCCACACGGTGAACCCCTCTCTACTATAAAATTAAGC 4

RESULT 7
AI886860
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AI886860 465 bp mRNA linear EST 21-DEC-1999
Wk29e07.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2413764 3'
similar to contains Alu repetitive element; mRNA sequence.
AI886860
AI886860.1 GI:5592024
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 465)
NCI/NINDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BTGP), Tumor Gene Index
Unpublished (1998)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1260 Std Error: 0.00

FEATURES
source

Seq primer: -40UP from Gibco
High quality sequence stop: 449.
Location/Qualifiers
1..465
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2413764"
/clone_lib="NCI_CGAP_Brn25"
/tissue_type="anaplastic oligodendroglioma"
/lab_host="DH10B"
/note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAGTGGGACGCCGCCATGAGTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."
BASE COUNT 142 a 112 c 99 g 112 t
ORIGIN

Query Match 5.7%; Score 57; DB 9; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.5e-16;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 169 GAGACCGCTGGCCACACGGTGAACCCCTCTCTACTATAAAATTAAGC 225
Db 269 GAGACCGCTGGCCACACGGTGAACCCCTCTCTACTATAAAATTAAGC 325

RESULT 8
BM353489/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BM353489 476 bp mRNA linear EST 07-JAN-2002
I955g11.x1 HR85 islet Homo sapiens cDNA 3', mRNA sequence.
BM353489
BM353489.1 GI:18085847
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 476)
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Lemishka, J., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas
M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T.,
Jackson, Y., and Bowers, Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohpc.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohpc.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40UP from Gibco
High quality sequence stop: 451.
Location/Qualifiers
1..476
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HR85 islet"
/tissue_type="purified pancreatic islet"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:

FEATURES
source

Location/Qualifiers
1..476
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HR85 islet"
/tissue_type="purified pancreatic islet"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:

NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
 Size-selected on agarose gel. Average insert size -1kb. 5'
 XhoI site was destroyed after directional cloning.
 Amplified once. Contact information: Hiroshi Inoue, MD,
 Metabolism Div. (Alan Permutt Lab), Washington University
 School of Medicine, Box 8127, 660 South Euclid Ave., St.
 Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
 314-362-1916, Fax: 314-747-2692.*

BASE COUNT 110 a 108 c 108 g 150 t

ORIGIN

Query Match 5.7%; Score 57; DB 13; Length 476;
 Best Local Similarity 100.0%; Pred. No. 2.5e-16;
 Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 GAGACAGCGCTGGCCACACACCGGTGAACCCCATCTCTACTTAAATAAATAAATAGC 225
 |||||
 Db 209 GAGACAGCGCTGGCCACACACCGGTGAACCCCATCTCTACTTAAATAAATAAATAGC 153

RESULT 9
 AA402129/c
 LOCUS 558 bp mRNA linear EST 09-NOV-1997
 DEFINITION zU55a08.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
 IMAGE:741878 5' similar to contains Alu repetitive element; contains
 element PRT5 repetitive element ; mRNA sequence.

ACCESSION AA402129
 VERSION AA402129.1 GI:2056094
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 558)
 Hillier, L., Allien, M., Bowles, L., Dubaque, J., Geisel, G., Jost, S.,
 Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin
 J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
 White, Y., Wylie, T., Waterston, R. and Wilson, R.

WASHU-NCI Human EST Project

Unpublished (1997)

Contact: Willson RK

Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 641 Std Error: 0.00

Seq primer: -28ml3 rev2 ET from Amersham

High quality sequence stop: 483.

FEATURES

source

1. 558
 Location/Qualifiers

/organism="Homo sapiens"

/db_xref="GDB:5942115"

/db_xref="taxon:9606"

/clone_lib="Soares ovary tumor NbHOT"

/sex="Female"

/tissue_type="Ovarian tumor"

/lab_host="DH10B (ampicillin resistant)"

/note="Organ: ovary; Vector: pT73D (Pharmacia) with a
 modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5'
 TGTTCACATCTGAAGTGGCGCGCGGTTTTTTTTTTTTTTT 3']
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT73 vector
 (Pharmacia). Library constructed by Bento Soares and
 M.Fatima Bonaldo."

BASE COUNT 130 a 118 c 122 g 188 t

ORIGIN

Query Match 5.7%; Score 57; DB 9; Length 558;
 Best Local Similarity 100.0%; Pred. No. 2.4e-16;
 Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 GAGACAGCGCTGGCCACACACCGGTGAACCCCATCTCTACTTAAATAAATAAATAGC 225
 |||||
 Db 286 GAGACAGCGCTGGCCACACACCGGTGAACCCCATCTCTACTTAAATAAATAAATAGC 230

RESULT 10

BQ646204/c

LOCUS 872 bp mRNA linear EST 15-JUL-2002

DEFINITION AGENCOURT_828784 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6297925
 5', mRNA sequence.

ACCESSION BQ646204

VERSION BQ646204.1 GI:21770376

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 872)

NIH-MGC http://mhc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: CGAP (Stanford)

cDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: Agencourt Bioscience Corporation

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LICM2508 row: c column: 14

High quality sequence stop: 657.

Location/Qualifiers

1. 872

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="IMAGE:6297925"

/clone_lib="NIH_MGC_100"

/tissue_type="hepatocellular carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: liver; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library."

BASE COUNT 190 a 195 c 222 g 265 t

ORIGIN

Query Match 5.7%; Score 57; DB 14; Length 872;
 Best Local Similarity 100.0%; Pred. No. 2.4e-16;
 Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 GAGACAGCGCTGGCCACACACCGGTGAACCCCATCTCTACTTAAATAAATAAATAGC 225
 |||||
 Db 677 GAGACAGCGCTGGCCACACACCGGTGAACCCCATCTCTACTTAAATAAATAAATAGC 621

RESULT 11

BQ648991/c

LOCUS 928 bp mRNA linear EST 15-JUL-2002

DEFINITION AGENCOURT_8418992 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6284304

5', mRNA sequence.

ACCESSION BQ648991

VERSION BQ648991.1 GI:21773163

KEYWORDS EST.

```

SOURCE
ORGANISM      human
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS      NIH-MGC http://mgs.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: CGAP (Stanford)
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: L1CM2481 row: 1 column: 01
              High quality sequence stop: 575.
FEATURES
source
1. .928
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /clone="IMAGE:6284304"
   /clone_lib="NIH_MGC_100"
   /tissue_type="hepatocellular carcinoma, cell line"
   /lab_host="DH10B (phage-resistant)"
   /note="Organ: liver; Vector: pOTB7; Site_1: XhoI; Site_2:
   EcoRI; cDNA made by oligo-dT priming. Directionally cloned
   into EcoRI/XhoI sites using the following 5' adaptor:
   GGCACGAG(G). Size-selected >500bp for average insert size
   1.8kb. Library constructed by Ling Hong in the laboratory
   of Gerald M. Rubin (University of California, Berkeley)
   using ZAP-cDNA synthesis kit (Stratagene) and Superscript
   II RT (Life Technologies). Note: this is a NIH_MGC
   Library."
BASE COUNT    212 a 217 c 230 g 269 t
ORIGIN
Query Match    5.7%; Score 57; DB 14; Length 928;
Best Local Similarity 100.0%; Pred. No. 2.3e-16;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 169 GAGACACGCTGGCCACACGGTGAAACCCATCTCTACTTAAATAAAAAATTAGC 225
|||||
Db 678 GAGACACGCTGGCCACACGGTGAAACCCATCTCTACTTAAATAAAAAATTAGC 622
|||||
RESULT 12
AL554221
LOCUS      AL554221 LTI_NFL006_PL2 948 bp mRNA linear EST 16-FEB-2001
DEFINITION prime, mRNA sequence.
ACCESSION  AL554221
VERSION     AL554221.1 GI:12894792
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS    Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL    Unpublished (2001)
COMMENT    Contact: Genoscope
              Genoscope - Centre National de Sequencage
              BP 191 91006 Evry cedex - France
              Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1. .948
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /clone="CSODI082YG14"

```

```

/clone_lib="LTI_NFL006_PL2"
/tissue_type="Placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT    284 a 231 c 210 g 214 t
ORIGIN
Query Match    5.7%; Score 57; DB 9; Length 948;
Best Local Similarity 100.0%; Pred. No. 2.3e-16;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 169 GAGACACGCTGGCCACACGGTGAAACCCATCTCTACTTAAATAAAAAATTAGC 225
|||||
Db 206 GAGACACGCTGGCCACACGGTGAAACCCATCTCTACTTAAATAAAAAATTAGC 262
|||||
RESULT 13
BM803038/c
LOCUS      BM803038 1086 bp mRNA linear EST 05-MAR-2002
DEFINITION AGENCOURT_6457358 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5560474
5', mRNA sequence.
ACCESSION  BM803038
VERSION     BM803038.1 GI:19119861
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS    NIH-MGC http://mgs.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Life Technologies, Inc.
              DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: L1AM12286 row: 1 column: 11
              High quality sequence stop: 606.
FEATURES
source
1. .1086
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /clone="IMAGE:5560474"
   /clone_lib="NIH_MGC_88"
   /tissue_type="Duodenal adenocarcinoma, cell line"
   /lab_host="PH10B (phage-resistant)"
   /note="Organ: small intestine; Vector: pCMV-SPORT6;
   Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
   oligo-dT primed. Average insert size 1.767 kb. Library
   enriched for full-length clones and constructed by Life
   Technologies. Note: this is a NIH_MGC Library."
BASE COUNT    203 a 328 c 234 g 321 t
ORIGIN
Query Match    5.7%; Score 57; DB 14; Length 1086;
Best Local Similarity 100.0%; Pred. No. 2.3e-16;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 169 GAGACACGCTGGCCACACGGTGAAACCCATCTCTACTTAAATAAAAAATTAGC 225
|||||

```

Db 627 GAGACCGCTGGCCACACAGGTGAACCCCTCTCTACTATAAAATTAAGATAGC 571

RESULT 14
AV718952/c

LOCUS AV718952 480 bp mRNA linear EST 16-OCT-2000

DEFINITION AV718952 GLC Homo sapiens cDNA clone GLCPC02 5', mRNA sequence.

ACCESSION AV718952

VERSION AV718952.1 GI:10816104

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Ou,J., Song,H., Cheng,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gu,Y., Chen,Z., and Han,Z.

TITLE Homo sapiens cDNA GLC clones

JOURNAL Unpublished (2000)

COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
1. 480
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GLCPC02"
/clone_lib="GLC"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 107 a 127 c 106 g 140 t

ORIGIN

Query Match 5.5%; Score 55; DB 10; Length 480;
Best Local Similarity 100.0%; Pred. No. 2.3e-15;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 232 GCGGTGGCTCACACCTGTAATCCAGCACTTGGAGCTGAGACGGTGGATCA 286
|||||

Db 265 GCGGTGGCTCACACCTGTAATCCAGCACTTGGAGCTGAGACGGTGGATCA 211
|||||

RESULT 15
AQ020028/c

LOCUS AQ020028 501 bp DNA linear GSS 09-JUN-1998

DEFINITION CIT-HSP-2312P17.TR CIT-HSP Homo sapiens genomic clone 2312P17, DNA sequence.

ACCESSION AQ020028

VERSION AQ020028.1 GI:3198764

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.

TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map Building (1998)

JOURNAL Unpublished (1998)

COMMENT Other GSSs: CIT-HSP-2312P17.TF
Contact: Mark Adams

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES

Location/Qualifiers
1. 501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2312P17"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2: HindIII"

BASE COUNT 117 a 124 c 109 g 151 t

ORIGIN

Query Match 5.5%; Score 55; DB 17; Length 501;
Best Local Similarity 100.0%; Pred. No. 2.2e-15;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 355 AATTAGCCAGGTGGTGGCACACCGCTGTAGTCCAGCTACTTGGAGGCTGAG 409
|||||

Db 95 AATTAGCCAGGTGGTGGCACACCGCTGTAGTCCAGCTACTTGGAGGCTGAG 41
|||||

RESULT 16

BG912932
LOCUS BG912932 629 bp mRNA linear EST 05-JUN-2001

DEFINITION 602807332P1 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4939639 5', mRNA sequence.

ACCESSION

VERSION BG912932.1 GI:14293408

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

COMMENT

Email: cgabbs@email.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10876 row: p column: 08
High quality sequence stop: 627.

FEATURES

Location/Qualifiers
1. 629
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4939639"
/clone_lib="NCI_CGAP_Brn67"
/tissue_type="anaplastic oligodendroglioma with lp/19q loss"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: Brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.3 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

```

BASE COUNT      189 a   133 c   150 g   156 t   1 others
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 55; DB 13; Length 629;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 340 TACTAAATAACAAAATACCCAGGTGGTGGGCACACGCCCTGTAGTCCAGCT 394
      |||||||
Db 447 TACTAAATAACAAAATACCCAGGTGGTGGGCACACGCCCTGTAGTCCAGCT 501
      |||||||

RESULT 17
AG158222/c
LOCUS
DEFINITION
AUI58222 PLACE1 Homo sapiens cDNA clone PLACE1011576 3', mRNA
sequence.
ACCESSION
AUI58222
VERSION
AUI58222.1 GI:11019743
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 479)
AUTHORS
Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Ozawa,M., Nakamura,Y., Nagai,T., Sugano
,S. and Isogai,T.Y.
TITLE
HRI human cDNA project (Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S.,
Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Ozawa,M., Nakamura
,Y., Nagai,T., Sugano,S., Isogai,T.)
JOURNAL
Unpublished (2000)
COMMENT
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
source
1..479
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="PLACE1"
/tissue_type="placenta"
/note="Vector: PME18SFL3"
BASE COUNT      108 a   120 c    99 g   143 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 54; DB 9; Length 479;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GGTGCGGTGCTACACCTTAATCCAGCACTTTGGGAGCGTGAGACGGGTG 281
      |||||||
Db 280 GGTGCGGTGCTACACCTTAATCCAGCACTTTGGGAGCGTGAGACGGGTG 227
      |||||||

RESULT 18
AG168389
LOCUS
DEFINITION
AG168389 Pan troglodytes DNA, clone: RP43-036L23.T7, genomic survey
sequence.
ACCESSION
AG168389
VERSION
AG168389.1 GI:16698067
KEYWORDS
GSS.
SOURCE
Pan troglodytes male lymphocytes DNA, clone_lib:RPCI-43 Chimpanzee
Male BAC Library clone:RP43-036L23.T7.

```

```

ORGANISM
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1
AUTHORS
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE
BAC end sequences of Library RPCI-43
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 727)
AUTHORS
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE
Direct Submission
JOURNAL
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:champs@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT
Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
end was generated during the Rad process and may have higher chance
of clone tracking errors.
PRIMERS
Sequencing: T7
LIBRARY
Vector : pBACe3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI
FEATURES
Location/Qualifiers
source
1..727
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="RP43-036L23.T7"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RPCI-43 Chimpanzee Male BAC Library"
BASE COUNT      258 a   167 c   150 g   150 t   2 others
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 54; DB 17; Length 727;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 344 AAAAATCAAAAATTAGCCAGGTGTGTGGCAGCGCTGTAGTCCAGCTACT 397
      |||||||
Db 241 AAAAATCAAAAATTAGCCAGGTGTGTGGCAGCGCTGTAGTCCAGCTACT 294
      |||||||

RESULT 19
BM477806/c
LOCUS
DEFINITION
AGENCOURT_6485596 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5555085
5', mRNA sequence.
ACCESSION
BM477806
VERSION
BM477806.1 GI:18526848
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1103)
AUTHORS
NIH-MGC http://mgi.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12274 row: k column: 22
High quality sequence stop: 636.

```

FEATURES
source

Location/Qualifiers
1. .1103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5555085"
/clone_lib="NIH_MGC_71"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: PCMV-SPORT6; Site:1; NotI;
Site:2; SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb. " 1 others
BASE COUNT 308 a 204 c 208 g 382 t
ORIGIN

Query Match 5.4%; Score 54; DB 13; Length 1103;
Best Local Similarity 100.0%; Pred. No. 6.3e-15;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 234 GGTGGCTCACACCTGTAAATCCAGCACTTTGGAGGCTGAGACGGGTGGATCAC 287
|||||
Db 74 GGTGGCTCACACCTGTAAATCCAGCACTTTGGAGGCTGAGACGGGTGGATCAC 21

RESULT 20
A2276522
LOCUS 280 bp DNA linear GSS 26-JUL-2000
DEFINITION RPCI-23-134121.TV RPCI-23 Mus musculus genomic clone RPCI-23-134121
REFERENCE DNA sequence.
ACCESSION A2276522
VERSION A2276522
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 280)
Zhao, S., Niemman, W., Feldblyum, T., Malek, J., Shatsman, S., Akhmet
B., Levins, M., Megann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P.
and Fraser, C.M.
TITLE Mouse PAC End Sequences from Library RPCI-23
JOURNAL Unpublished (1999)
COMMENT Other_GSSs: RPCI-23-134121.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 0208
Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html
Plate: 134 row: 1 column: 21
Seq primer: T7
Class: BAC ends.
FEATURES
source

Location/Qualifiers
1. .280
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-134121"
/clone_lib="RPCI-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site:1;
ECORI; Site:2; EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBACe3.6 vector at the
ECORI sites. The ligation products were transformed into

BASE COUNT 104 a 45 c 32 g 99 t
ORIGIN

Query Match 5.3%; Score 53; DB 17; Length 280;
Best Local Similarity 100.0%; Pred. No. 2.1e-14;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 509 AATAATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATCT 561
|||||
Db 97 AATAATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATCT 149

RESULT 21
AA552202/c
LOCUS 320 bp mRNA linear EST 17-MAR-1999
DEFINITION nj89c04.s1 NCI_CGAP_Prl1 Homo sapiens cDNA clone IMAGE:99654
similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION AA552202
VERSION AA552202
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 320)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@rs-remail.nih.gov

Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Rodrigo F.
Chuaqui, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbrp/image/image.html
Insert length: 475 Std Error: 0.00
Seq primer: -40ml3 fwd. RT from Amersham
High quality sequence stop: 287.
FEATURES
source

Location/Qualifiers
1. .320
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:99654"
/clone_lib="NCI_CGAP_Prl1"
/sex="male"
/tissue_type="normal prostatic epithelial cells"
/lab_host="DH10B"
/note="Organ: prostate; Vector: pAMP10; mRNA made from
normal prostatic epithelial cells, cDNA made by oligo-dT
priming. Non-directionally cloned. Size-selected on
agarose gel, average insert size 600 bp. Library made by
D. Krizman, NIH."
BASE COUNT 67 a 88 c 74 g 91 t
ORIGIN

Query Match 5.3%; Score 53; DB 9; Length 320;
Best Local Similarity 100.0%; Pred. No. 2.1e-14;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 322 ATGTGAACCCAGCTCTTACTATAATAATAATAATAATAATAATAATAGCCAGTGTGGTGC 374
|||||
Db 189 ATGTGAACCCAGCTCTTACTATAATAATAATAATAATAATAATAATAGCCAGTGTGGTGC 137

RESULT 22
BF149427/c
LOCUS 329 bp mRNA linear EST 26-OCT-2000
DEFINITION 18_19 Human Epidermal Keratinocyte Subtraction Library- Upregulated

Transcripts Homo sapiens cDNA similar to gi 2178259, mRNA sequence.
ACCESSION BF149427
VERSION BF149427.1 GI:11030822
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 329)
AUTHORS Schlager,J.J., Benjamin,H.R., Ali,K., Levine,C.F., Dodds,A.D., Avery,D.P. and Clark,J.H.
TITLE Human Epidermal Keratinocyte Subtraction Library- Upregulated Transcripts
JOURNAL Unpublished (2000)
COMMENT Contact: Schlager,J.J.
Molecular Toxicology
United States Army Medical Research Institute of Chemical Defense (USAMRICD)
3100 Ricketts Point Road; ATTN: MCMR-UV-PA; APG-EA, MD 21010-5400, USA
Tel: 410 436 1940
Fax: 410 436 1960
Email: John.Schlager@AMEDD.ARMY.MIL.
Location/Qualifiers
1..329
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human Epidermal Keratinocyte Subtraction Library- Upregulated Transcripts"
/sex="female"
/tissue_type="skin"
/cell_type="primary keratinocyte"
/cell_line="Clonetics 4075"
/dev_stage="adult"
/lab_host="E. coli (ampicillin-resistant)"
/note="Vector: pT-Adv; Library preparation: Sequence isolated using PCR-select (Clontech) subtraction library construction (Diatchenko et al., (1996) Proc. Natl. Acad. Sci. USA 93: 6025-6030) after sulfur mustard exposure of primary human adult epidermal keratinocytes. Subtraction library cloned for random sequence selection into pT-Adv vector."
BASE COUNT 60 a 81 c 77 g 111 t
ORIGIN
Query Match 5.3%; Score 53; DB 12; Length 329;
Best Local Similarity 100.0%; Pred. No. 2.le-14;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 169 GAGACAGCTGGCCACACGCGTGAACCCCATCTCTACTAAATAAAAT 221
|||||
DB 304 GAGACAGCTGGCCACACGCGTGAACCCCATCTCTACTAAATAAAAT 252
|||||
RESULT 23
ACQ125843/c
LOCUS 351 bp DNA linear GSS 23-SEP-1998
DEFINITION HS 2168.B2.D11.MF CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate-2168 Col-22 Row=H, DNA sequence.
ACCESSION ACQ125843
VERSION ACQ125843.1 GI:3503009
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 351)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

99380589
MEDLINE
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 2168 row: H column: 22
Class: BAC ends
High quality sequence stop: 351.
Location/Qualifiers
1..351
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Plate-2168 Col=22 Row=H"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"
BASE COUNT 67 a 103 c 85 g 96 t
ORIGIN
Query Match 5.3%; Score 53; DB 17; Length 351;
Best Local Similarity 100.0%; Pred. No. 2.le-14;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 234 GGTGCTCACACCTGTAAATCCAGCACTTTGGAGGCTGAGACGGTGGATCA 286
|||||
DB 253 GGTGCTCACACCTGTAAATCCAGCACTTTGGAGGCTGAGACGGTGGATCA 201
|||||
RESULT 24
AA411448
LOCUS 352 bp mRNA linear EST 17-MAY-1997
DEFINITION zV30e05.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone IMAGE:755168 3' similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION AA411448
VERSION AA411448.1 GI:2068980
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 352)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,S., Jost,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wyllie,T., Waterston,R. and Wilson,R.
TITLE WashU-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -41ml3 fwd. ET from Amersham
High quality sequence stop: 311.
Location/Qualifiers
1..352
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:755168"
/clone_lib="Soares ovary tumor NbHOT"
/sex="Female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: ovary; Vector: pT73D (Pharmacia) with a

modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGGGTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo. "

BASE COUNT 117 a 76 c 70 g 89 t

ORIGIN

Query Match 5.3%; Score 53; DB 9; Length 352;
Best Local Similarity 100.0%; Pred. No. 2.le-14;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 GAGACAGCTGGCCACACACGGTGAACCCCATCTCTACTATAAATAAAAAAT 221
|||||
Db 274 GAGACAGCTGGCCACACACGGTGAACCCCATCTCTACTATAAATAAAAAAT 326
|||||

RESULT 25
AA455483 491 bp mRNA linear EST 06-JUN-1997
LOCUS
DEFINITION
xx7606.sl Soares ovary tumor NbHOT Homo sapiens cDNA clone
IMAGE:809722 3' similar to contains Alu repetitive element;; mRNA
sequence.

ACCESSION
AA455483
VERSION
AA455483.1 GI:2178259
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.,
Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,
T., Waterston,R. and Wilson,R.
WashU-Merck EST Project 1997
Unpublished (1997)
Contact: Willson RK

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 417.

FEATURES
Location/Qualifiers
1..491
/organism="Homo sapiens"
/db_xref="GDB:6039414"
/db_xref="taxon:9606"
/clone="IMAGE:809722"
/clone_lib="Soares ovary tumor NbHOT"
/sex="Female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: ovary; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGGGTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo. "

BASE COUNT 165 a 111 c 108 g 107 t

ORIGIN

Query Match 5.3%; Score 53; DB 9; Length 491;

Best Local Similarity 100.0%; Pred. No. 2e-14;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 GAGACAGCTGGCCACACACGGTGAACCCCATCTCTACTATAAATAAAAAAT 221
|||||
Db 267 GAGACAGCTGGCCACACACGGTGAACCCCATCTCTACTATAAATAAAAAAT 319
|||||

RESULT 26
AW973181/c 577 bp mRNA linear EST 01-JUN-2000
LOCUS
DEFINITION
EST385279 WAGE rsequences, MAGM Homo sapiens cDNA, mRNA sequence.
ACCESSION
AW973181
VERSION
AW973181.1 GI:8163027
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt
I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
Quackenbush,J.
Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
Unpublished (2000)
Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tigr.org
Plate: 315

Seq primer: Forward.
Location/Qualifiers
1..577
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="WAGE rsequences, MAGM"
/note="Vector: pBluescriptSkm"
BASE COUNT 146 a 148 c 125 g 158 t
ORIGIN

Query Match 5.3%; Score 53; DB 10; Length 577;
Best Local Similarity 100.0%; Pred. No. 2e-14;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 322 ATGTGTGAACACGCTCTCTACTATAAATAACAAAATTAGCCAGGTGTGTGGC 374
|||||
Db 305 ATGTGTGAACACGCTCTCTACTATAAATAACAAAATTAGCCAGGTGTGTGGC 253
|||||

RESULT 27
AG012178/c 691 bp DNA linear GSS 14-APR-1999
LOCUS
DEFINITION
Homo sapiens genomic DNA, 21q region, clone: T1957BG11, genomic
survey sequence.
ACCESSION
AG012178 AG005159
VERSION
AG012178.1 GI:3410839
KEYWORDS
GSS.
SOURCE
Homo sapiens DNA, clone:T1957BG11.
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE
Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
JOURNAL
Homo sapiens genomic DNA, chromosome 21q
Published Only in Database (1998)
REFERENCE
2 (bases 1 to 691)
AUTHORS
Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE
Direct Submission
JOURNAL
Submitted (09-AUG-1998) Masahira Hattori, Kitasato University,
Department of Science, JST Sequencing Laboratory, Kitasato 1-15-1,

```

Sagamihara 228, Japan (E-mail:hattori@hgc.ims.u-tokyo.ac.jp,
Tel:0427-78-9732; Fax:0427-78-9561)
On Feb 5, 1999 this sequence version replaced gi:2865117.
Ag005159: Submitted (10-Feb-1998).
FEATURES
    source
        1..691
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="21"
            /map="21q"
            /clone="T1957BG11"
BASE COUNT    170 a   160 c   150 g   209 t       2 others
ORIGIN
Query Match          5.3%; Score 53; DB 17; Length 691;
Best Local Similarity 100.0%; Pred. No. 2e-14;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GTGGCTCACACCTGTAATCCAGCAGCTTTGGAGGCTGAGACGGGTGGATCAC 287
|||||
DB 384 GTGGCTCACACCTGTAATCCAGCAGCTTTGGAGGCTGAGACGGGTGGATCAC 332
|||||

RESULT 28
AG012177/c
LOCUS
DEFINITION Homo sapiens genomic DNA, 21q region, clone: T1957BG11, genomic
survey sequence.
ACCESSION AG012177 AG005158
VERSION AG012177.1 GI:3410838
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone:T1957BG11.
ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
    Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
    1 (bases 1 to 693)
    TITLE Homo sapiens genomic DNA, chromosome 21q
    JOURNAL Published Only in Database (1998)
AUTHORS
    Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
DIRECT SUBMISSION
    Submitted (09-AUG-1998) Masahira Hattori, Kitasato University,
    Department of Science, JST Sequencing Laboratory; Kitasato 1-15-1,
    Sagamihara 228, Japan (E-mail:hattori@hgc.ims.u-tokyo.ac.jp,
    Tel:0427-78-9732; Fax:0427-78-9561)
    On Feb 5, 1999 this sequence version replaced gi:2865116.
    AG005158: Submitted (10-Feb-1998).
FEATURES
    source
        1..693
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="21"
            /map="21q"
            /clone="T1957BG11"
BASE COUNT    199 a   166 c   144 g   179 t       5 others
ORIGIN
Query Match          5.3%; Score 53; DB 17; Length 693;
Best Local Similarity 100.0%; Pred. No. 2e-14;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GTGGCTCACACCTGTAATCCAGCAGCTTTGGAGGCTGAGACGGGTGGATCAC 287
|||||
DB 275 GTGGCTCACACCTGTAATCCAGCAGCTTTGGAGGCTGAGACGGGTGGATCAC 223
|||||

RESULT 29
AG012190
LOCUS
DEFINITION Homo sapiens genomic DNA, 21q region, clone: T1957BG11, genomic
survey sequence.
ACCESSION AG012190 AG005171

```

```

AG012190.1 GI:3410851
GSS.
Homo sapiens DNA, clone:T1957BG11.
Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
    Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
    1 (bases 1 to 702)
    TITLE Homo sapiens genomic DNA, chromosome 21q
    JOURNAL Published Only in Database (1998)
AUTHORS
    Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
DIRECT SUBMISSION
    Submitted (09-AUG-1998) Masahira Hattori, Kitasato University,
    Department of Science, JST Sequencing Laboratory; Kitasato 1-15-1,
    Sagamihara 228, Japan (E-mail:hattori@hgc.ims.u-tokyo.ac.jp,
    Tel:0427-78-9732; Fax:0427-78-9561)
    On Feb 5, 1999 this sequence version replaced gi:2865127.
    AG005171: Submitted (10-Feb-1998).
FEATURES
    source
        1..702
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="21"
            /map="21q"
            /clone="T1957BG11"
BASE COUNT    239 a   150 c   163 g   146 t       4 others
ORIGIN
Query Match          5.3%; Score 53; DB 17; Length 702;
Best Local Similarity 100.0%; Pred. No. 2e-14;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GTGGCTCACACCTGTAATCCAGCAGCTTTGGAGGCTGAGACGGGTGGATCAC 287
|||||
DB 86 GTGGCTCACACCTGTAATCCAGCAGCTTTGGAGGCTGAGACGGGTGGATCAC 138
|||||

RESULT 30
AG012191
LOCUS
DEFINITION Homo sapiens genomic DNA, 21q region, clone: T1957BG11, genomic
survey sequence.
ACCESSION AG012191 AG005172
VERSION AG012191.1 GI:3410852
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone:T1957BG11.
ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
    Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
    1 (bases 1 to 705)
    TITLE Homo sapiens genomic DNA, chromosome 21q
    JOURNAL Published Only in Database (1998)
AUTHORS
    Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
DIRECT SUBMISSION
    Submitted (09-AUG-1998) Masahira Hattori, Kitasato University,
    Department of Science, JST Sequencing Laboratory; Kitasato 1-15-1,
    Sagamihara 228, Japan (E-mail:hattori@hgc.ims.u-tokyo.ac.jp,
    Tel:0427-78-9732; Fax:0427-78-9561)
    On Feb 5, 1999 this sequence version replaced gi:2865128.
    AG005172: Submitted (10-Feb-1998).
FEATURES
    source
        1..705
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="21"
            /map="21q"
            /clone="T1957BG11"
BASE COUNT    217 a   143 c   172 g   166 t       7 others
ORIGIN

```

Query Match 5.3%; Score 53; DB 17; Length 705;
Best Local Similarity 100.0%; Pred. No. 2e-14;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GTGGCTCACCTGTATCCAGACATTTGGAGGCTGAGACGGGTGATC 287
|||||
Db 342 GTGGCTCACCTGTATCCAGACATTTGGAGGCTGAGACGGGTGATC 394
|||||

RESULT 31
BJ339324

LOCUS BJ339324 293 bp mRNA linear EST 07-MAR-2002
DEFINITION BJ339324 Dictyostelium discoidium cDNA library, AF Dictyostelium
discoidium cDNA clone dda64n19 5', mRNA sequence.

ACCESSION BJ339324
VERSION BJ339324.1 GI:19247686
KEYWORDS EST.
SOURCE Dictyostelium discoidium.
ORGANISM Dictyostelium discoidium
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE 1 (bases 1 to 293)
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
TITLE Full length cDNA of Dictyostelium discoidium at the aggregation
stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..293
/organism="Dictyostelium discoidium"
/strain="AX4"
/db_xref="taxon:44689"
/clone_lib="Dictyostelium discoidium cDNA library, AF"
/dev_stage="Aggregation stage"
BASE COUNT 177 a 14 c 18 g 84 t
ORIGIN

Query Match 5.2%; Score 52; DB 13; Length 293;
Best Local Similarity 100.0%; Pred. No. 6.4e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 507 AGAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 558
|||||
Db 66 AGAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 117
|||||

RESULT 32
BJ396501

LOCUS BJ396501 293 bp mRNA linear EST 08-MAR-2002
DEFINITION BJ396501 Dictyostelium discoidium cDNA library, SF Dictyostelium
discoidium cDNA clone dds42102 5', mRNA sequence.

ACCESSION BJ396501
VERSION BJ396501.1 GI:19307587
KEYWORDS EST.
SOURCE Dictyostelium discoidium.
ORGANISM Dictyostelium discoidium
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE 1 (bases 1 to 293)
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
TITLE Full length cDNA of Dictyostelium discoidium at the slug stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856

Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..293
/organism="Dictyostelium discoidium"
/strain="AX4"
/db_xref="taxon:44689"
/clone_lib="Dictyostelium discoidium cDNA library, SF"
/sex="mat A"
/dev_stage="Slug stage"
BASE COUNT 177 a 14 c 18 g 84 t
ORIGIN

Query Match 5.2%; Score 52; DB 13; Length 293;
Best Local Similarity 100.0%; Pred. No. 6.4e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 507 AGAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 558
|||||
Db 66 AGAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 117
|||||

RESULT 33
BJ366477

LOCUS BJ366477 303 bp mRNA linear EST 08-MAR-2002
DEFINITION BJ366477 Dictyostelium discoidium cDNA library, CF Dictyostelium
discoidium cDNA clone ddc39102 5', mRNA sequence.

ACCESSION BJ366477
VERSION BJ366477.1 GI:19275779
KEYWORDS EST.
SOURCE Dictyostelium discoidium.
ORGANISM Dictyostelium discoidium
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE 1 (bases 1 to 303)
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
TITLE Full length cDNA of Dictyostelium discoidium at the culmination
stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..303
/organism="Dictyostelium discoidium"
/strain="AX4"
/db_xref="taxon:44689"
/clone_lib="Dictyostelium discoidium cDNA library, CF"
/sex="mat A"
/dev_stage="Culmination stage"
BASE COUNT 179 a 14 c 19 g 91 t
ORIGIN

Query Match 5.2%; Score 52; DB 13; Length 303;
Best Local Similarity 100.0%; Pred. No. 6.4e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 507 AGAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 558
|||||
Db 76 AGAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 127
|||||

RESULT 34
BJ394919

LOCUS BJ394919 305 bp mRNA linear EST 08-MAR-2002
DEFINITION BJ394919 Dictyostelium discoidium cDNA library, SF Dictyostelium
discoidium cDNA clone dds37b01 5', mRNA sequence.

ACCESSION BJ394919

VERSION BJ394919.1 GI:19306005
 KEYWORDS EST.
 SOURCE Dictyostelium discoideum.
 ORGANISM Dictyostelium discoideum
 Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
 REFERENCE 1 (bases 1 to 305)
 AUTHORS Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
 TITLE Full length cDNA of Dictyostelium discoideum at the slug stage
 JOURNAL Unpublished (2002)
 COMMENT Contact: Tadasi Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp.

FEATURES
 source
 1..305
 Location/Qualifiers
 /organism="Dictyostelium discoideum"
 /strain="AX4"
 /db_xref="taxon:44689"
 /clone="dds37b01"
 /clone_lib="Dictyostelium discoideum cDNA library, SF"
 /sex="mat A"
 /dev_stage="Slug stage"
 /dev_stage="Slug stage" 19 g 92 t

BASE COUNT 180 a 14 c 19 g 92 t
 ORIGIN

Query Match 5.2%; Score 52; DB 13; Length 305;
 Best Local Similarity 100.0%; Pred. No. 6.4e-14;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 507 AGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 558
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 78 AGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 129

RESULT 35
 AQ350115 521 bp DNA linear GSS 07-MAY-1999
 LOCUS RPC111-110C18.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-110C18,
 DEFINITION DNA sequence.
 ACCESSION AQ350115
 VERSION AQ350115.1 GI:4177450
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 521)
 AUTHORS Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter, J.C.

TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
 JOURNAL Map Building
 COMMENT Unpublished (1997)

Other_GSSs: RPC111-110C18.TV
 Contact: Shaying Zhao, William Nierman, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208

Email: hbe@tigr.org
 Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
 Research Genetics (info@resgen.com). BAC end search page:
 http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 source
 1..521
 Location/Qualifiers

/organism="Homo sapiens"
 /db_xref="GDB:7541921"
 /db_xref="taxon:9606"
 /clone="RPCI-11-110C18"
 /clone_lib="RPCI-11"
 /sex="Male"
 /cell_type="Lymphocytes"
 /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
 RPC111 Human Male BAC Library"
 BASE COUNT 206 a 103 c 97 g 115 t
 ORIGIN

Query Match 5.2%; Score 52; DB 17; Length 521;
 Best Local Similarity 100.0%; Pred. No. 6.1e-14;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 334 CGTCTCTACTAAATACAAAATAGCCAGGTGTGGTGGCACACGCTGTA 385
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 392 CGTCTCTACTAAATACAAAATAGCCAGGTGTGGTGGCACACGCTGTA 443

RESULT 36
 AQ376489/c 596 bp DNA linear GSS 20-MAY-1999
 LOCUS RPC111-149C17.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-149C17,
 DEFINITION DNA sequence.
 ACCESSION AQ376489.1 GI:4347512
 VERSION AQ376489
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 596)
 AUTHORS Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter, J.C.
 TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
 JOURNAL Map Building
 COMMENT Unpublished (1997)

Other_GSSs: RPC111-149C17.TV
 Contact: Shaying Zhao, William Nierman, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbe@tigr.org
 Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
 Research Genetics (info@resgen.com). BAC end search page:
 http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 source
 Location/Qualifiers
 1..596
 /organism="Homo sapiens"
 /db_xref="GDB:7556896"
 /db_xref="taxon:9606"
 /clone="RPCI-11-149C17"
 /clone_lib="RPCI-11"
 /sex="Male"
 /cell_type="Lymphocytes"
 /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
 RPC111 Human Male BAC Library"
 BASE COUNT 177 a 88 c 71 g 260 t
 ORIGIN

Query Match 5.2%; Score 52; DB 17; Length 596;
 Best Local Similarity 100.0%; Pred. No. 6e-14;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 507 AGAAATATAATATAATATAATATAATATAATATAATATAATATAATATAAT 558
Db 540 AGAAATATAATATAATATAATATAATATAATATAATATAATATAATATAAT 489

RESULT 37
AI909139
LOCUS AI909139 601 bp mRNA linear EST 30-MAR-2000
DEFINITION IL-BT200-010499-002 BT200 Homo sapiens cDNA, mRNA sequence.
ACCESSION AI909139
VERSION AI909139.1 GI:6499819
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 601)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/seq/gethtml.pl?tl=il&t2=il-BT200-002.html&
st3=010499&t4=1)
Seq primer: puc 18 forward.
FEATURES
source
Location/Qualifiers
1..601
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BT200"
/sex="female"
/dev_stage="Adult"
/notes="Organ: breast; Vector: puc18; Site:1; SmaI; Site:2;
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 171 a 112 c 171 g 135 t 12 others
ORIGIN
Query Match 5.2%; Score 52; DB 9; Length 601;
Best Local Similarity 100.0%; Pred. No. 6e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 236 TGGCTCACACCTGTAATCCAGCACTTTGGAGGCTTGACAGGGTGGATCAC 287
Db 227 TGGCTCACACCTGTAATCCAGCACTTTGGAGGCTTGACAGGGTGGATCAC 278

RESULT 38
BI861844
LOCUS BI861844 617 bp mRNA linear EST 10-OCT-2001
DEFINITION 60338872F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5398121 5',
mRNA sequence.
ACCESSION BI861844
VERSION BI861844.1 GI:16002591

```

```

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 617)
NIH-MGC http://mgc.nci.nih.gov/.
NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)
UNPUBLISHED (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12014 row: o column: 18
High quality sequence stop: 578.
Location/Qualifiers
1..617
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5398121"
/clone_lib="NIH_MGC_87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DHI0B (phage-resistant)"
/notes="Organ: breast; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SmaI; Cloned unidirectionally; oligo-dr primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 193 a 147 c 161 g 116 t
ORIGIN
Query Match 5.2%; Score 52; DB 13; Length 617;
Best Local Similarity 100.0%; Pred. No. 6e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 CGCTCTACTAAAAATACAAAATAGCCAGGTGGTGGCACACGCCTGTA 385
Db 89 CGCTCTACTAAAAATACAAAATAGCCAGGTGGTGGCACACGCCTGTA 140

RESULT 39
AZ955691/c
LOCUS AZ955691 643 bp DNA linear GSS 27-APR-2001
DEFINITION 2M0221F23R Mouse 10Kb plasmid UUGC2M library Mus musculus genomic
clone UUGC2M0221F23 R, DNA sequence.
ACCESSION AZ955691
VERSION AZ955691.1 GI:13826918
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 643)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
UNPUBLISHED (2000)
CONTACT: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu

```

Insert Length: 10000 Std Error: 0.00
 Plate: 0221 row: F column: 23
 Seq primer: CACACAGGAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 643.
 Location/Qualifiers
 1. .643

FEATURES

source
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone_lib="M0221F23"
 /clone_lib="Mouse 10kb plasmid UUGC2M library"
 /sex="Female"
 /lab_host="E. coli strain XL10-Gold, Tl-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
 BASE COUNT 218 a 101 c 120 g 204 t
 ORIGIN

Query Match 5.2%; Score 52; DB 17; Length 643;
 Best Local Similarity 100.0%; Pred. No. 6e-14;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 510 AATAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATTTCT 561
 Db 435 AATAAAATAAATAAATAAATAAATAAATAAATAAATAAATTTCT 384

RESULT 40
 AQ740087
 LOCUS
 DEFINITION HS_5503_A2_H09_T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=1079 Col=18 Row=0, DNA sequence.
 AQ740087
 VERSION
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 AUTHORS Mahairas G.G., Wallace J.C., Smith K., Swartzell S., Holzman T., Keller A., Shaker R., Furlong J., Young J., Zhao S., Adams M.D. and Hood L.
 TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPCI-11. For BAC

library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Resear h Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu
 Plate: 1079 row: O column: 18
 Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 865.
 Location/Qualifiers
 1. .865

FEATURES

source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="Plate=1079 Col=18 Row=0"
 /clone_lib="RPCI-11 Human Male BAC Library"
 /sex="Male"
 /note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACE3.6 vector at EcoRI sites"
 BASE COUNT 336 a 178 c 153 g 198 t
 ORIGIN

Query Match 5.2%; Score 52; DB 17; Length 865;
 Best Local Similarity 100.0%; Pred. No. 5.8e-14;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 334 CCGTCTCTACTAAAATACAAAATAGCCAGGTGGTGGGCACGCGCTGTA 385
 Db 411 CCGTCTCTACTAAAATACAAAATAGCCAGGTGGTGGGCACGCGCTGTA 462

RESULT 41
 BE883679
 LOCUS
 DEFINITION 601507363F1 NTH_MGC_71 Homo sapiens cDNA clone IMAGE:3908942 5', mRNA sequence.
 BE883679
 ACCESSION
 VERSION BE883679.1 GI:10332455
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 962)
 AUTHORS NTH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Plate: LLAM9721 row: n column: 15
 High quality sequence stop: 569.
 Location/Qualifiers
 1. .962

FEATURES

source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3908942"
 /clone_lib="NTH_MGC_71"
 /tissue_type="Leiomysarcoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 2.1 kb."
 BASE COUNT 298 a 216 c 283 g 164 t
 ORIGIN 1 others

Query Match 5.2%; Score 52; DB 12; Length 962;
Best Local Similarity 100.0%; Pred. No. 5.8e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 GGTCTCTACTTAAATACAAAATAGCCAGGTGGTGGCAGCCCTGTGA 385
|||||
Db 124 GGTCTCTACTTAAATACAAAATAGCCAGGTGGTGGCAGCCCTGTGA 175

RESULT 42
AW795350/c

LOCUS AW795350 306 bp mRNA linear EST 16-MAY-2000
DEFINITION PM0-UM0018-130500-003-g05 UM0018 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW795350
VERSION AW795350.1 GI:7847220
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 306)

AUTHORS Dias Neto,E., Garcia Correa,R., Verjovsky-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Garvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Bruenstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=622=PM0-UM0018-130
500-003-g05&t3=2000-05-13&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 306.

FEATURES

source
1..306
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UM0018"
/dev_stage="Adult"

/note="Organ: uterus; Vector: puc18; Site.1: SmaI; Site.2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT 70 a 65 c 74 g 97 t
ORIGIN

Query Match 5.1%; Score 51; DB 10; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.9e-13;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 GAGACAGCCTGGCCACACGGTGAACCCCATCTCTACTATAAATAAAAA 219
|||||

Db 89 GAGACAGCCTGGCCACACGGTGAACCCCATCTCTACTATAAATAAAAA 39
|||||

RESULT 43
AZ342219

LOCUS AZ342219 569 bp DNA linear GSS 29-SEP-2000
DEFINITION 1M0075K08F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
Clone UUGCLM0075K08 F, DNA sequence.

ACCESSION AZ342219
VERSION AZ342219.1 GI:10419239
KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 569)

AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Lonacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D.,Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL Unpublished (2000)

COMMENT

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: adunn@genetics.utah.edu

Insert Length: 1000 Std Error: 0.00

Plate: 0075 row: K column: 08

Seq primer: CGTTGTAACGACGCGCCAGT

Class: plasmid ends

High quality sequence stop: 569.

FEATURES

Location/Qualifiers
1..569
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCLM0075K08"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/sex="Male"

/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gil147321141gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 202 a 120 c 90 g 157 t
ORIGIN

Query Match 5.1%; Score 51; DB 17; Length 569;
Best Local Similarity 100.0%; Pred. No. 1.8e-13;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 508 GAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 558
|||||

Db 233 GAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 283
|||||

RESULT 44
AZ118261/c

LOCUS AZ118261 588 bp DNA linear GSS 12-MAY-2000
DEFINITION RPCI-23-450D11.TJ RPCI-23 Mus musculus genomic clone RPCI-23-450D11
ACCESSION AZ118261
VERSION AZ118261.1 GI:7782025
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 588)
AUTHORS Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
, B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
and Fraser,C.M.
TITLE Mouse BAC End Sequences from Library RPCI-23
JOURNAL Unpublished (1999)
COMMENT Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 450 row: D column: 11
Seq primer: SP6
Class: BAC ends
FEATURES
source
Location/Qualifiers
1..588
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-450D11"
/clone_lib="RPCI-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBACe3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT 158 a 118 c 107 g 205 t
ORIGIN
Query Match 5.1%; Score 51; DB 17; Length 588;
Best Local Similarity 100.0%; Pred. No. 1.8e-13;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 509 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATT 559
|||||
Db 472 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATT 422
RESULT 45
AQ932513/c
LOCUS AQ932513 606 bp DNA linear GSS 21-DEC-1999
DEFINITION RPCI-23-286D16.TJ RPCI-23 Mus musculus genomic clone RPCI-23-286D16
ACCESSION AQ932513
VERSION AQ932513.1 GI:6621527
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 606)
AUTHORS Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
, B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
and Fraser,C.M.
TITLE Mouse BAC End Sequences from Library RPCI-23
JOURNAL Unpublished (1999)
COMMENT Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 286 row: D column: 16
Seq primer: SP6
Class: BAC ends
FEATURES
source
Location/Qualifiers
1..606
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-286D16"
/clone_lib="RPCI-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBACe3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT 167 a 123 c 116 g 199 t
ORIGIN
Query Match 5.1%; Score 51; DB 17; Length 606;
Best Local Similarity 100.0%; Pred. No. 1.8e-13;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 509 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATT 559
|||||
Db 461 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATT 411
Search completed: November 24, 2002, 12:00:49
Job time : 2204 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2002, 09:16:57 ; Search time 293 Seconds
(without alignments)
7686.001 Million cell updates/sec

Title: US-09-784-423-32
Perfect score: 1000
Sequence: 1 GGTGTGACCTTATCCTCTCT.....CAGATGATGACCGCGTGC 1000

Scoring table: OLIGO_NUC

Gapop 60.0 , Capext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Length	ID	Description
1	1000	100.0	1000	20	AA227732 Human DNA marker c
2	72	7.2	52616	22	AAK70459 Human immune/haema
3	72	7.2	52616	22	AAK78930 Human immune/haema
4	58	5.8	11627	22	AAI07363 Human reproductive
5	58	5.8	11627	23	AB198832 Human testicular a
6	57	5.7	29329	22	ABA18026 Human nervous syst
7	57	5.7	29329	22	ABA20511 Human nervous syst
8	57	5.7	29329	22	AAK70791 Human immune/haema
9	57	5.7	29329	22	AAK78512 Human immune/haema

c	10	56	5.6	122888	24	ABK83569	Human cDNA differe
c	11	55	5.5	476	22	AAF65641	Novel human polynu
c	12	54	5.4	479	22	AAH13151	Human cDNA clone (
c	13	54	5.4	593	24	AAH089372	Human prostate exp
c	14	54	5.4	2298	22	AAH16608	Human cDNA sequenc
c	15	53	5.3	307	22	AAK32680	Human genomic DNA
c	16	53	5.3	642	22	AAK45882	Human full-length
c	17	52	5.2	700	22	AAH92081	Human inflammatory
c	18	52	5.2	700	22	AAH92081	Human inflammatory
c	19	52	5.2	2701	22	AAK37013	Human musculoskele
c	20	52	5.2	5533	22	AAK74315	Human immune/haema
c	21	52	5.2	5533	22	AAK74317	Human immune/haema
c	22	52	5.2	5536	22	AAK74316	Human immune/haema
c	23	52	5.2	5537	22	AAK74318	Human immune/haema
c	24	52	5.2	12244	22	AAK69991	Human digestive sy
c	25	52	5.2	12392	22	ABA15896	Human nervous syst
c	26	52	5.2	1503900	22	AAK95240	Human neuregulin-1
c	27	52	5.2	1503900	22	AAK95240	Human neuregulin-1
c	28	51	5.1	6523	24	ABN80120	Human chemically m
c	29	51	5.1	43411	24	ABQ88169	Human osteoblast d
c	30	51	5.1	227968	24	ABK83497	Human cDNA differe
c	31	50	5.0	276	22	AAK36586	Human cardiovascular
c	32	50	5.0	276	22	AAK75295	Human immune/haema
c	33	50	5.0	276	22	AAK75296	Human immune/haema
c	34	50	5.0	475	23	ABV54360	Human prostate exp
c	35	50	5.0	578	23	ABV55359	Human prostate exp
c	36	50	5.0	737	22	AAI94916	Human neuroblastom
c	37	50	5.0	4243	22	ABA16762	Human nervous syst
c	38	50	5.0	5283	22	AAK74563	Human immune/haema
c	39	50	5.0	5572	22	AAK74320	Human immune/haema
c	40	50	5.0	5573	22	AAK74321	Human immune/haema
c	41	50	5.0	5649	22	AAK46383	Tumour suppressor
c	42	50	5.0	5649	22	AAK46384	Tumour suppressor
c	43	50	5.0	5649	24	ABK40007	Human chemically p
c	44	50	5.0	5649	24	ABK40008	Human chemically p
c	45	50	5.0	5649	24	ABK32848	Human immune syste

ALIGNMENTS

RESULT 1
AA227732
ID AA227732 standard; DNA; 1000 BP.
XX
AC AA227732;
XX
DT 23-DEC-1999 (first entry)
XX
DE Human DNA marker clone S132.
XX
KW Tandem repeat sequence; DNA isolation; intermediate tandem repeat;
KW ITR sequence; pentanucleotide tandem repeat; stutter artifact;
KW DNA typing; DNA profiling; linkage analysis; criminal justice;
KW paternity testing; animal lineage analysis; microsatellite loci;
KW polymorphism detection; ds.
XX
OS Homo sapiens.
XX
PN W09940194-A1.
XX
PD 12-AUG-1999.
XX
PF 04-FEB-1999; 99WO-US02345.
XX
PR 04-FEB-1998; 98US-0018584.
XX
PA (PROM-) PROMEGA CORP.
XX
PI Schumm JW, Bacher JW;
XX
DR WPI; 1999-590696/50.
XX

Isolating DNA containing intermediate tandem repeat sequences, useful in DNA profiling -
 Claim 22; Page 78; 11pp; English.

This sequence represents a human DNA marker clone used in the method of the invention. The method is for isolating a fragment of DNA containing an intermediate tandem repeat (ITR) sequence using hybridization selection, and comprises: (a) providing several DNA fragments, at least one of which contains an ITR sequence, a region of the DNA fragment which contains at least one repeat unit consisting of a sequence of five, six or seven bases repeated in tandem at least two times; (b) providing a stationary support having at least one oligonucleotide associated with it, where the oligonucleotide includes a sequence of nucleotides which is complementary to a portion of the ITR sequence; and (c) combining the DNA fragments with the support under conditions where the DNA fragments including the DNA fragment containing the ITR sequence hybridize to the support. The method is particularly used to isolate DNA containing pentanucleotide tandem repeat sequences as well as to detect target ITR DNA sequences having a low incidence of stutter artifacts (no more than 2.4%). The method is useful in DNA profiling for linkage analysis, criminal justice, paternity testing and other forensic and medical uses. DNA typing is also useful for confirming the lineage of horses, dogs and other prize animals. The invention overcomes problems related to the use of microsatellite loci in DNA profiling. The method can detect polymorphisms with a low incidence of stutter artifacts, which has previously been a problem in interpreting allelic content of loci. The development of markers based on larger repeat units, enables easier separation of the fragments on electrophoretic gels. This allows the simultaneous analysis of more loci.

Sequence 1000 BP; 302 A; 240 C; 231 G; 226 T; 1 other;

Query Match 100.0%; Score 1000; DB 20; Length 1000;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGTGACCTTATCCTCTGTAACCTCAGTTTCCTCATCCGTTAAATGAAGAGCTGTAG 60
 DB 1 GGTGTGACCTTATCCTCTGTAACCTCAGTTTCCTCATCCGTTAAATGAAGAGCTGTAG 60
 QY 61 ATTGTGTAATAAATAAATGAATAGGCTAGCGGCTGCCTCAGCTGTAATCCCA 120
 DB 61 ATTGTGTAATAAATAAATGAATAGGCTAGCGGCTGCCTCAGCTGTAATCCCA 120
 QY 121 GCACCTTAGAGGTGACAGAGGTGGATCAGTTCAGTTCAGGAGTTTGAGACCAAGCTG 180
 DB 121 GCACCTTAGAGGTGACAGAGGTGGATCAGTTCAGGAGTTTGAGACCAAGCTG 180
 QY 181 GCCAACACGGTGAACCCCATCTCTACTATAAATAAATAAATAAATAAATAAATAAATAA 240
 DB 181 GCCAACACGGTGAACCCCATCTCTACTATAAATAAATAAATAAATAAATAAATAAATAA 240
 QY 241 CACACCTGTAATCCAGCAGCTTTGGAGGCTGAGACGGTGGATCAGCTGAAGTCAGGAG 300
 DB 241 CACACCTGTAATCCAGCAGCTTTGGAGGCTGAGACGGTGGATCAGCTGAAGTCAGGAG 300
 QY 301 TTCAAGGCCAGCTGGGCAACATGGTGAACACAGCTCTCTACTATAAATAAATAAATAA 360
 DB 301 TTCAAGGCCAGCTGGGCAACATGGTGAACACAGCTCTCTACTATAAATAAATAAATAA 360
 QY 361 CCAGGTGTGGTGGCAGACAGCTGTAGTCCAGCTACTTTGGAGGCTGAGGCGGAAGATC 420
 DB 361 CCAGGTGTGGTGGCAGACAGCTGTAGTCCAGCTACTTTGGAGGCTGAGGCGGAAGATC 420
 QY 421 GCTTGAACCCAGTAGGAGAGGTTCAGTTCAGGCGGAGATAGAGTCACTGCCAGGCC 480
 DB 421 GCTTGAACCCAGTAGGAGAGGTTCAGTTCAGGCGGAGATAGAGTCACTGCCAGGCC 480
 QY 481 TGGGTGACAGACAGAGCTCCCTCTCAGAAATAAATAAATAAATAAATAAATAAATAA 540
 DB 481 TGGGTGACAGACAGAGCTCCCTCTCAGAAATAAATAAATAAATAAATAAATAAATAA 540

QY 541 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 600
 DB 541 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 600
 QY 601 AGTAATAGCTATCAATATATCCACCCCTACCACTGTCTGTAATAATTAGTTTCTTTTGTG 660
 DB 601 AGTAATAGCTATCAATATATCCACCCCTACCACTGTCTGTAATAATTAGTTTCTTTTGTG 660
 QY 661 ACCCCCCATTAAGGAGAGTAATCTCACCAGTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
 DB 661 ACCCCCCATTAAGGAGAGTAATCTCACCAGTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
 QY 721 GGCACATAGTGGGTCTCAGTGAACATGCTGAGTGAAGAGCAAAATGCAAGGAATCTCC 780
 DB 721 GGCACATAGTGGGTCTCAGTGAACATGCTGAGTGAAGAGCAAAATGCAAGGAATCTCC 780
 QY 781 AGGCCATCTGGAGAGCCCTCCAGCGGGTGTAGTTGCGGAAACTCATAGTCTCTCTCAAT 840
 DB 781 AGGCCATCTGGAGAGCCCTCCAGCGGGTGTAGTTGCGGAAACTCATAGTCTCTCTCAAT 840
 QY 841 GGCCTACTGAAAGGTAGAGAGTTCTGGTCCCACTCGGCACCCCACTCTCTGACTCAC 900
 DB 841 GGCCTACTGAAAGGTAGAGAGTTCTGGTCCCACTCGGCACCCCACTCTCTGACTCAC 900
 QY 901 TGTGTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 960
 DB 901 TGTGTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 960
 QY 961 CAGGACTGCAAGGAGCCGACAGATGATGACCGCGTGC 1000
 DB 961 CAGGACTGCAAGGAGCCGACAGATGATGACCGCGTGC 1000

RESULT 2
 AAK70459
 ID AAK70459 standard; DNA; 52616 BP.
 XX
 AC AAK70459;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:25271.
 XX
 KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer; cytostatic; gene therapy; vaccine; metastasis; ds.
 OS Homo sapiens.
 XX
 PN W0200157182-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01354.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-02245119.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226686.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0234998.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.

PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0251799.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-483426/52.
 XX
 DR
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Disclosure; SEQ ID NO 25271, 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK67694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK62169

CC represent sequences used in the exemplification of the present invention.

XX Sequence 52616 BP; 13218 A; 13917 C; 13903 G; 11678 T; 0 other;

SQ Query Match 7.2%; Score 72; DB 22; Length 52616;

Best Local Similarity 100.0%; Pred. No. 3 4e-24;

Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 CGTCTCTACTTAAATACAAAAATTAGCCAGGTGGTGGCACACGCTGTAGTCCACG 393

Db 23964 CGTCTCTACTTAAATACAAAAATTAGCCAGGTGGTGGCACACGCTGTAGTCCACG 24023

QY 394 TACTTGGGAGGC 405

Db 24024 TACTTGGGAGGC 24035

RESULT 3

AAK78930

ID AAK78930 standard; DNA: 52616 BP.

XX AC AAK78930;

XX DT 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:33742.

XX Human; immune; haematopoietic; Immune/haematopoietic antigen; cancer;

XX Cytostatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX WO200157182-A2.

XX PD 09-AUG-2001.

XX FF 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179055.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 11-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217496.

XX 14-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220963.

XX 26-JUL-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224518.

XX 14-AUG-2000; 2000US-0224519.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225214.

XX 14-AUG-2000; 2000US-0225266.

XX 14-AUG-2000; 2000US-0225267.

XX 14-AUG-2000; 2000US-0225268.

XX 14-AUG-2000; 2000US-0225270.

XX 14-AUG-2000; 2000US-0225447.

XX 14-AUG-2000; 2000US-0225757.

XX 14-AUG-2000; 2000US-0225758.

XX 14-AUG-2000; 2000US-0225759.

XX 18-AUG-2000; 2000US-0226279.

XX 22-AUG-2000; 2000US-0226681.

XX 22-AUG-2000; 2000US-0226688.

XX 22-AUG-2000; 2000US-0227182.

XX 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.

PR 08-SEP-2000; 2000US-0232080.

PR 08-SEP-2000; 2000US-0232081.

PR 12-SEP-2000; 2000US-0231968.

PR 14-SEP-2000; 2000US-0232397.

PR 14-SEP-2000; 2000US-0232398.

PR 14-SEP-2000; 2000US-0232399.

PR 14-SEP-2000; 2000US-0232400.

PR 14-SEP-2000; 2000US-0232401.

PR 14-SEP-2000; 2000US-0233063.

PR 14-SEP-2000; 2000US-0233064.

PR 14-SEP-2000; 2000US-0233065.

PR 21-SEP-2000; 2000US-0234223.

PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234597.

PR 25-SEP-2000; 2000US-0234998.

PR 26-SEP-2000; 2000US-0235484.

PR 27-SEP-2000; 2000US-0235834.

PR 27-SEP-2000; 2000US-0235836.

PR 29-SEP-2000; 2000US-0236327.

PR 29-SEP-2000; 2000US-0236367.

PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236369.

PR 29-SEP-2000; 2000US-0236370.

PR 02-OCT-2000; 2000US-0236802.

PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.

PR 02-OCT-2000; 2000US-0237039.

PR 13-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239937.

PR 20-OCT-2000; 2000US-0240960.

PR 20-OCT-2000; 2000US-0241221.

PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241786.

PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241808.

PR 20-OCT-2000; 2000US-0241809.

PR 20-OCT-2000; 2000US-0241826.

PR 01-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0246474.

PR 08-NOV-2000; 2000US-0246475.

PR 08-NOV-2000; 2000US-0246476.

PR 08-NOV-2000; 2000US-0246477.

PR 08-NOV-2000; 2000US-0246478.

PR 08-NOV-2000; 2000US-0246523.

PR 08-NOV-2000; 2000US-0246524.

PR 08-NOV-2000; 2000US-0246525.

PR 08-NOV-2000; 2000US-0246526.

PR 08-NOV-2000; 2000US-0246527.

PR 08-NOV-2000; 2000US-0246528.

PR 08-NOV-2000; 2000US-0246532.

PR 08-NOV-2000; 2000US-0246609.

PR 08-NOV-2000; 2000US-0246610.

PR 08-NOV-2000; 2000US-0246611.

PR 17-NOV-2000; 2000US-0249207.

PR 17-NOV-2000; 2000US-0249208.

PR 17-NOV-2000; 2000US-0249209.

PR 17-NOV-2000; 2000US-0249210.

PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WFL; 2001-483426/52.
DR
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis
XX
PS Disclosure; SEQ ID NO 33742; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 52616 BP; 13218 A; 13817 C; 13903 G; 11678 T; 0 other;

Query Match 7.2%; Score 72; DB 22; Length 52616;
Best Local Similarity 100.0%; Pred. No. 3.4e-24;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 CTTCTCTACTAAAAATACAAAAATAGCCAGGTGGTGGCACACGCTGTAGTCCCGC 333
Db 23964 CGTCTCTACTAAAAATACAAAAATAGCCAGGTGGTGGCACACGCTGTAGTCCCGC 333
QY 394 TACTTGGGAGGC 405
Db 24024 TACTTGGGAGGC 24035

RESULT 4
AAL07363/C
ID AAL07363 standard; DNA; 11627 BP.
XX
AC AAL07363;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen DNA SEQ ID NO: 10051.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200155320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01339.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0188874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 05-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.

PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 21-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 25-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239945.
 PR 20-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0244674.
 PR 08-NOV-2000; 2000US-0244675.
 PR 08-NOV-2000; 2000US-0244676.
 PR 08-NOV-2000; 2000US-0244677.
 PR 08-NOV-2000; 2000US-0244678.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251889.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0254097.
 PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-483232/52.

DR Nucleic acids encoding 973 human testicular antigen polypeptides,
 PI useful for preventing, diagnosing and/or treating testicular cancer -

XX Disclosure; SEQ ID NO 3484; 766pp; English.

CC The present invention provides the protein and coding sequences of 973
 CC human testicular antigens, and fragments of their genomic sequences. The
 CC sequences can be used in the treatment of cardiovascular, urinary system,
 CC reproductive system, immune, respiratory, neurological and
 CC gastrointestinal disorders, infections, and particularly cancer,
 CC especially testicular cancers. The present sequence is a DNA encoding a
 CC protein fragment of the invention.

XX SQ Sequence 11627 BP; 3510 A; 2177 C; 2243 G; 3697 T; 0 other;
 Query Match 5.8%; Score 58; DB 23; Length 11627;
 Best Local Similarity 100.0%; Pred. No. 1.5e-17;
 Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 340 TACTAAATAACAAAATAGCCAGGTGTGGTGGCACGCCGTAGTCCAGCTACT 397
 Db 10819 TACTAAATAACAAAATAGCCAGGTGTGGTGGCACGCCGTAGTCCAGCTACT 10762
 RESULT 6
 ABAL8026/c
 ID ABAL8026 standard; DNA; 29329 BP.
 XX AC ABA18026;
 XX DT 23-JAN-2002 (first entry)
 XX DE Human nervous system related polynucleotide SEQ ID NO 10357.
 XX KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
 KW antiparkinsonian; antischoling; antianemic; antiarthritic; cancer;
 KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
 KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
 KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
 OS Homo sapiens.
 XX PN W0200159063-A2.
 XX PD 16-AUG-2001.
 XX PF 17-JAN-2001; 2001WO-US01334.
 XX PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226688.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 06-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 21-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 20-OCT-2000; 2000US-0242221.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.

PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249219.
 PR 17-NOV-2000; 2000US-0249220.
 PR 17-NOV-2000; 2000US-0249221.
 PR 17-NOV-2000; 2000US-0249222.
 PR 17-NOV-2000; 2000US-0249223.
 PR 17-NOV-2000; 2000US-0249224.
 PR 17-NOV-2000; 2000US-0249225.
 PR 17-NOV-2000; 2000US-0249226.
 PR 17-NOV-2000; 2000US-0249227.
 PR 17-NOV-2000; 2000US-0249228.
 PR 17-NOV-2000; 2000US-0249229.
 PR 17-NOV-2000; 2000US-0249230.
 PR 01-DEC-2000; 2000US-0250391.
 PR 01-DEC-2000; 2000US-0251160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0251989.
 PR 06-DEC-2000; 2000US-0251479.
 PR 06-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 PR Rosen CA, Barash SC, Ruben SM;
 PI WPI; 2001-541565/60.
 DR
 XX
 PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating nervous system
 PT cancers and metastases
 XX
 PS Disclosure; SEQ ID NO 10357; 1701pp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABAI1004-ABA21334) and proteins
 CC (ABAI14678-ABAI18001) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 29329 Bp; 6864 A; 6756 C; 7225 G; 8484 T; 0 other;
 Query Match 5.7%; Score 57; DB 22; Length 29329;
 Best Local Similarity 100.0%; Pred. No. 4.5e-17;
 Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 235 GTGGCTCACCTGATCCAGCAGCTTTGGAGGCTGAGCGGTGATCCTGA 291
 Db 11642 GTGGCTCACCTGATCCAGCAGCTTTGGAGGCTGAGCGGTGATCCTGA 11586
 RESULT 7
 ABA20511/c

ID ABA20511 standard; DNA; 29329 Bp.
 XX
 AC ABA20511;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Human nervous system related polynucleotide SEQ ID NO 12842.
 XX
 KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulvaric;
 KW antiparkinsonian; antiskinning; antianemic; antiarthritic; cancer;
 KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
 KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
 KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200159063-A2.
 XX
 PD 16-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01334.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0215647.
 PR 07-JUL-2000; 2000US-0215680.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 14-AUG-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 18-AUG-2000; 2000US-0225759.
 PR 22-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 23-AUG-2000; 2000US-0227182.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 08-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.

PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233084.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239955.
PR 20-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.

PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
XX useful for preventing, diagnosing and/or treating nervous system
XX cancers and metastases -
XX
XX Disclosure; SEQ ID NO 12842; 1701pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AB11004-ABA21534) and proteins
XX (AB14678-AB18001) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful
XX in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
XX and ovarian cancer and other cancers of the adrenal gland, bone, bone
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
XX (b) immune disorders e.g. Addison's disease, allergies, autoimmune
XX haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative
XX colitis; (c) cardiovascular disorders such as myocardial ischaemia;
XX (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
XX and parasitic infections.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 29329 BP; 6864 A; 6756 C; 7225 G; 8484 T; 0 other;

Query Match 5.7%; Score 57; DB 22; Length 29329;
Best Local Similarity 100.0%; Pred. No. 4.5e-17;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GTGGCTCACACCTGTATCCACACACTTGGAGCGCTGAGACGGTGGATCACTGA 291
|||||
Db 11642 GTGGCTCACACCTGTATCCACACACTTGGAGCGCTGAGACGGTGGATCACTGA 11586

RESULT 8
AAK70791/c
ID AAK70791 standard; DNA; 29329 BP.
XX
XX AAK70791;
AC
XX
XX
DT 06-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:25603.
DE
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
KW
XX
OS Homo sapiens.
XX
PN WO200157182-A2.

[illegible]

```
DR WPI; 2001-483426/52.
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis
XX Disclosure; SEQ ID NO 25603; 3071pp + Sequence Listing; English.
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK62170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 29329 BP; 6864 A; 6756 C; 7225 G; 8484 T; 0 other;

Query Match          5.7%; Score 57; DB 22; Length 29329;
Best Local Similarity 100.0%; Pred. No. 4.5e-17;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GTGGCTCACACCTGTATCCACAGCTTTGGAGGCTGACACCGGTGGATCACTGA 291
      |||||||
Db 11642 GTGGCTCACACCTGTATCCACAGCTTTGGAGGCTGACACCGGTGGATCACTGA 11586

RESULT 9
AAK78512/c
ID AAK78512 standard; DNA; 29329 BP.
AC AAK78512;
XX
XX 07-NOV-2001 (first entry)
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:33324.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216647.
XX 07-JUL-2000; 2000US-0216880.
XX 11-JUL-2000; 2000US-0217487.
XX 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224516.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236357.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
```

PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 05-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX
XX Disclosure; SEQ ID NO 33324; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,

CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 29329 BP; 6864 A; 6756 C; 7225 G; 8484 T; 0 other;

Query Match 5.7%; Score 57; DB 22; Length 29329;
Best Local Similarity 100.0%; Pred. No. 4.5e-17;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GTGGCTCACACCTGTATCCAGACACTTGGGAGGCTGAGACGGTGGATCACTGA 291
|||||
Db 11642 GTGGCTCACACCTGTATCCAGACACTTGGGAGGCTGAGACGGTGGATCACTGA 11586
|||||

RESULT 10
ABK83569/C
ID ABK83569 standard; cDNA; 122888 BP.
XX
XX AC ABK83569;
XX
DT 14-AUG-2002 (first entry)
XX
XX Human cDNA differentially expressed in granulocytic cells #140.
XX
XX Human; ss; granulocytic cell; DNA chip; bacterial infection;
XX viral infection; parasitic infection; protozoal infection;
XX fungal infection; sterile inflammatory disease; psoriasis;
XX rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
XX cardiac reperfusion injury; renal reperfusion injury; ARDS;
XX adult respiratory distress syndrome; inflammatory bowel disease;
XX Crohn's disease; ulcerative colitis; periodontal disease;
XX granulocyte activation; chronic inflammation; allergy.
XX
XX Homo sapiens.
XX
XX WO200228999-A2.
XX
PD 11-APR-2002.
XX
XX 03-OCT-2001; 2001WO-US30821.
XX
XX PF
XX PR 03-OCT-2000; 2000US-237189P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX WPI; 2002-435328/46.
XX
XX Detecting granulocyte activation by detecting differential expression
XX of genes associated with granulocyte activation, which serves as
XX diagnostic markers that is useful for monitoring disease states and
XX drug toxicity -
XX
XX Claim 1; SEQ ID No 140; 114pp; English.
XX
XX The invention relates to detecting (M1) granulocyte (GC) activation
XX (GCA), by detecting the level of expression of gene(s) (Gs) identified by
XX DNA chip analysis as given in the specification, and comparing
XX the expression level to an expression level in an unactivated
XX GC, where differential expression of Gs is indicative of GCA.
XX Also included are modulating (M2) GA by contacting GC with an agent
XX that alters the expression of at least one gene in Gs; (2) screening (M3)
XX for an agent capable of modulating GCA or an inflammation (especially
XX chronic) in a tissue, an allergic response in a subject, exposure of a
XX subject to a pathogen or sterile inflammatory disease using the
XX gene expression profile; (3) detecting (M4) an inflammation (especially
XX chronic) in a tissue, an allergic response in a subject, exposure of a
XX subject to a pathogen or sterile inflammatory disease, by detecting the
XX level of expression in a sample of the tissue of gene(s) from Gs, where

the level of expression of the gene is indicative of inflammation;
 (4) treating (M5) an inflammation (especially chronic) or in a tissue,
 an allergic response in a subject, exposure of a subject to a pathogen
 or sterile inflammatory disease, by contacting a tissue having
 inflammation with an agent that modulates the expression of gene(s)
 from GS in the tissue. M1 is useful for detecting GCA; M2 is useful for
 modulating GA; M3 is useful for screening an agent capable of modulating
 GCA preferably in an inflammation in a tissue; M4 is useful for
 detecting an inflammation (especially chronic) in a tissue, an allergic
 response in a subject, exposure of a subject to a pathogen or sterile
 inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
 glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
 reperfusion injury, ARDS, adult respiratory distress syndrome,
 inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 periodontal disease; also bacterial infection, viral infection,
 parasitic infection, protozoal infection, fungal infection and M5 is
 useful for treating one of the above conditions. The present
 sequence represents a gene differentially expressed in granulocytes.
 Note: The sequence data for this patent did not form part
 of the printed specification, but was obtained in electronic
 format directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences.

Sequence 122888 BP; 28761 A; 33410 C; 31919 G; 28798 T; 0 other;

Query Match 5.6%; Score 56; DB 24; Length 122888;
 Best Local Similarity 100.0%; Pred. No. 1.3e-16;
 Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 232 GCGGTGGCTCACCTGTATCCAGCACTTTGGAGGCTGAGCGGTGGATCAC 287
 |||||
 Db 55450 GCGGTGGCTCACCTGTATCCAGCACTTTGGAGGCTGAGCGGTGGATCAC 55395

RESULT 11

AAAF65641
 ID AAF65641 standard; cDNA; 476 BP.

AC AAF65641;

DT 09-APR-2001 (first entry)

DE Novel human polynucleotide, SEQ ID NO: 1397.

Human; cytostatic; gene therapy; colon cancer; prostate cancer;
 breast cancer; lung cancer; cancer detection; ss.

OS Homo sapiens.

PN WC200102568-A2.

PD 11-JAN-2001.

PF 30-JUN-2000; 2000WO-US18374.

PR 02-JUL-1999; 99US-0142310.

PR 02-JUL-1999; 99US-0142311.

PA (CHIR) CHIRON CORP.

(HYSE-) HYSEQ INC.

PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
 PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
 PI Crkenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
 PI Kita D, Garcia V, Jones LW, Strache-Crain B;

DR WPI; 2001-091805/10.

XX Library of polynucleotides for diagnosing a cancerous state of a
 PT mammalian cell and detecting cancer, particularly of the colon or
 PT prostate, comprises 3351 human polynucleotide sequences -

XX Claim 9; Page 744; 1046pp; English.

XX

CC The present sequence is one of 3351 sequences in a library of human
 CC polynucleotides. The library is used to detect differentially expressed
 CC genes correlated with a cancerous state of a mammalian cell and can
 CC detect colon, prostate, breast and lung cancer. The library can be used
 CC to produce probes for detection of mRNA and to produce additional copies
 CC of the polynucleotides. The probes can be used for chromosome mapping of
 CC the polynucleotide and for detection of transcription levels. Ribozymes
 CC or antisense oligonucleotides can be generated. The polynucleotides and
 CC their gene products are used as genetic or biochemical markers (e.g. in
 CC blood or tissues) that will detect the earliest changes along the
 CC carcinogenesis pathway and/or monitor the efficacy of therapies and
 CC preventive interventions. The polynucleotides, polypeptides and
 CC antibodies against them can be used in pharmaceutical compositions to
 CC treat the cancers and proliferative disorders such as neoplasia,
 CC dysplasia and hyperplasia.

SQ Sequence 476 BP; 124 A; 130 C; 120 G; 98 T; 4 other;

Query Match 5.5%; Score 55; DB 22; Length 476;
 Best Local Similarity 100.0%; Pred. No. 4.4e-16;
 Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 GAGACACGCTGGCCACACACGGTGAACCCCATCTCTACTAAAAATAATTA 223

|||||
 Db 305 GAGACACGCTGGCCACACACGGTGAACCCCATCTCTACTAAAAATAATTA 359

RESULT 12

AAH13151/C

ID AAH13151 standard; cDNA; 479 BP.

XX AAH13151;

DT 26-JUN-2001 (first entry)

DE Human cDNA clone (3'-primer) SEQ ID NO:9986.

Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

OS Homo sapiens.

PN EPI074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -

PS Claim 3; SEQ ID 9986; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesising 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH95993 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

SQ Sequence 479 BP; 108 A; 120 C; 99 G; 143 T; 9 other;

Query Match 5.4%; Score 54; DB 22; Length 479;
Best Local Similarity 100.0%; Pred. No. 1.3e-15;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GGGTGGGTGGCTCACACCTGTAATCCAGCAGCTTTGGAGGCTGAGACGGGTG 281
DB 280 GGGTGGGTGGCTCACACCTGTAATCCAGCAGCTTTGGAGGCTGAGACGGGTG 227
|||||

RESULT 13
ABQ89372

ID ABQ89372 standard; cDNA; 593 BP.

AC ABQ89372;

DT 27-SEP-2002 (first entry)

DE Human prostate expressed polynucleotide SEQ ID NO 628.

KW Human; prostate; cytostatic; tumour; cancer; vaccine; gene therapy;
KW gene; ss.

OS Homo sapiens.

PN W0200255700-A2.

XX 18-JUL-2002.

PF 07-DEC-2001; 2001WO-US47349.

PR 07-DEC-2000; 2000US-254648P.

XX 13-MAR-2001; 2001US-275688P.

PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.

PI Escobedo J, Garcia PD, Kassam A, Lamson G, Drmanac R;
PI Crkvenjakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones WL, Stache-Crain B, Scott EM;
XX WPI; 2002-557824/59.

XX New genes and gene products isolated from human prostate, useful for
PT treating or diagnosing tumor or cancer (e.g. prostate cancer or breast
PT cancer), or as vaccines for treating or preventing these diseases -

PS Claim 1; SEQ ID NO 628; 186pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide comprising any of
CC 1477 sequences or its fragment, degenerate variant, antisense or
CC complement. The polynucleotides and gene products are useful for treating
CC or diagnosing tumour or cancer (e.g. prostate cancer, breast cancer,
CC lung cancer or medullary carcinoma) in a subject (e.g. cattle, dogs,

CC cats, rabbits, horse or human). The polynucleotides and polypeptides are
CC also useful as vaccines for treating or preventing these diseases. The
CC polynucleotides are useful for gene therapy. The present sequence is that
CC of one of a group of polynucleotides (ABQ88745-ABQ90015) disclosed
CC electronically as sequences of the invention. However only 1271
CC polynucleotide sequences are given, whereas 1477 polynucleotides and 91
CC proteins are claimed.
CC Note: the sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequence.

SQ Sequence 593 BP; 186 A; 116 C; 141 G; 134 T; 16 other;

Query Match 5.4%; Score 54; DB 24; Length 593;
Best Local Similarity 100.0%; Pred. No. 1.3e-15;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GGGTGGGTGGCTCACACCTGTAATCCAGCAGCTTTGGAGGCTGAGACGGGTG 281
DB 94 GGGTGGGTGGCTCACACCTGTAATCCAGCAGCTTTGGAGGCTGAGACGGGTG 147
|||||

RESULT 14

AAH16608

ID AAH16608 standard; cDNA; 2298 BP.

XX AAH16608;

AC AAH16608;

DT 26-JUN-2001 (first entry)

DE Human cDNA sequence SEQ ID NO:15707.

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

OS Homo sapiens.

PN EPI074617-A2.

XX 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Mishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -

PS Claim 8; SEQ ID 15707; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH1742 represent human cDNA sequences; AAH92446 to
 CC AAH95893 represent human amino acid sequences; AAH92446 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX
 SQ Sequence 2298 BP; 764 A; 473 C; 481 G; 580 T; 0 other;

Query Match 5.4%; Score 54; DB 22; Length 2298;
 Best Local Similarity 100.0%; Pred. No. 1.3e-15;
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GGGTGGGGTGGCTACACCTGTATCCAGCACCTTGGGAGGCTGAGACGGGTG 281
 |||||||
 Db 2019 GGGTGGGGTGGCTACACCTGTATCCAGCACCTTGGGAGGCTGAGACGGGTG 2072

RESULT 15

AAH32680
 ID AAS32680 standard; DNA; 307 BP.

XX AAS32680;

DT 17-DEC-2001 (first entry)

DE Human genomic DNA for novel endocrine antigen, SEQ ID No 634.

KW Human; endocrine antigen; ds; cytostatic; antiinfertility; antidiabetic;
 KW thyroid-active; adrenal-active; androgenic; gastric; gene therapy;
 KW antisense-therapy; antibody; endocrine disorder; hormone imbalance;
 KW reproductive disorder; endocrine cancer; pancreatic disorder;
 KW diabetes mellitus; adrenal gland disorder; hirsutism; thyroid disorder;
 KW hyperthyroidism; hypothalamic disorder; vanishing testes syndrome.

OS Homo sapiens.

XX WO200155319-A2.

XX 02-AUG-2001.

PE 17-JAN-2001; 2001WO-US01335.

XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 15-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 14-JUL-2000; 2000US-0217496.
 PR 25-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 14-AUG-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 18-AUG-2000; 2000US-0225759.
 PR 22-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 23-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.

PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251038.
PR 06-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-457726/49.
XX
PT Isolated polypeptide for treating, preventing and prognosing disorders
PT related to the endocrine system including endocrine disorders,
PT reproductive disorders, and gastrointestinal disorders and also for
PT testing and detection e.g. diagnosis -
XX
PS Disclosure; SEQ ID NO 634; 558pp; English.
XX
CC The invention relates to cDNAs encoding novel human endocrine
CC antigens or a fragment having biological activity, a domain, an epitope,
CC full length protein, variant, allelic variant or a species homologue of
CC the cDNA/antigen. The DNAs and polypeptides are useful for preventing,
CC treating or ameliorating a medical condition when administered
CC (e.g. by gene therapy or antisense-therapy). Identifying mutations in
CC the genes coding for the antigens is useful for diagnosing a pathological
CC condition or a susceptibility to a pathological condition. The DNAs,
CC antigens and antibodies raised against the antigens useful for treating,
CC preventing and/or prognosing disorders related to the endocrine system
CC or hormone imbalance or reproductive disorders, cancers of endocrine
CC tissues, disorders of the pancreas (e.g. diabetes mellitus), the adrenal
CC glands (e.g. hirsutism), ovaries, the thyroid (e.g. hyperthyroidism), the
CC hypothalamus and testes (e.g. vanishing testes syndrome), many examples
CC of diseases and disorders are given in the specification. The present
CC sequence is genomic DNA fragment form a gene encoding an endocrine
CC antigen of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 307 BP; 91 A; 75 C; 93 G; 48 T; 0 other;
Query Match 5.3%; Score 53; DB 22; Length 307;
Best Local Similarity 100.0%; Pred. No. 4e-15;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 234 GGTGGCTCACACCTGTAAATCCAGCAGCTTTGGAGGCTGAGACGGTGGATCA 286
Db 14 GGTGGCTCACACCTGTAAATCCAGCAGCTTTGGAGGCTGAGACGGTGGATCA 66
RESULT 16
AAS44682/c
ID AAS44682 standard; DNA; 642 BP.
XX
AC AAS44682;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human full-length polynucleotide sequence #107.
XX
KW Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
KW nervous system disorder; inflammatory disorder; cell differentiation; ds;
KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;
KW cytostatic; antirheumatic; antiarthritic; vulnary; antinflammatory;
KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
KW neuroprotective; osteopathic; antidiabetic; antiallergic;
KW immunostimulant; analgesic; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200164834-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US04926.
XX
PR 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
PR 17-JUN-2000; 2000US-0597707.
PR 14-JUL-2000; 2000US-0616807.
PR 19-SEP-2000; 2000US-0664641.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
PI Drmanac R;
XX
DR WPI; 2001-589862/66.
DR P-PSDB; AAU27782.
XX
PT Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis, treatment of
PT cancer, neurological, inflammatory disorders and for use in arrays for
PT detection -
XX
PS Claim 1; SEQ ID No 107; 153pp; English.
XX
CC Sequences AAS44576-AAS44919 represent full-length polynucleotides and
CC contig polynucleotides encoding polypeptides of the invention. The DNA
CC and protein sequences are useful for the treatment, diagnosis and
CC prevention of various types of disorder in a mammalian subject such as a
CC human, dog, monkey, mouse, hamster or rat. The disorders include cancers
CC such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such
CC as multiple sclerosis, connective tissue disease, rheumatoid arthritis,
CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and

CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's
 CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
 CC bowel disease. The sequences exhibit activity relating to angiogenesis,
 CC cell proliferation, cell differentiation, stem cell growth factor,
 CC activin or inhibin. Therefore, they can be used to manipulate stem cells
 CC in culture to give rise to neuroepithelial cells that can be used to
 CC augment or replace cells damaged by illness, accidental damage or genetic
 CC disorders. The sequences may also be used for regeneration of bone,
 CC cartilage, tendons and ligaments and in tissue repair and burn healing.
 CC Note: Some sequences for this patent did not form part of the printed
 CC specification, but were obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 642 BP; 145 A; 130 C; 158 G; 209 T; 0 other;

Query Match 5.3%; Score 53; DB 22; Length 642;
 Best Local Similarity 100.0%; Pred. No. 3.9e-15;
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 GAGACAGCTGGCCACACGGTGAACCCATCTCTACTTAAATAAAAAAT 221
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 278 GAGACAGCTGGCCACACGGTGAACCCATCTCTACTTAAATAAAAAAT 225

RESULT 17

AAH92080
 ID AAH92080 standard; DNA; 700 BP.

XX
 AC AAH92080;

DT 09-OCT-2001 (first entry)

DE Human inflammatory bowel disease related gene fragment IGR2092a.

KN Human; inflammatory bowel disease; Crohn's disease; ulcerative colitis;
 KW single nucleotide polymorphism; SNP; chromosome 19p13; paternity test;
 KW chromosome 5q31-33; forensic test; gene therapy; ds.

XX Homo sapiens.

PN WO200142511-A2.

PD 14-JUN-2001.

PF 11-DEC-2000; 2000WO-US33632.

PR 10-DEC-1999; 99US-0170257.

PR 10-APR-2000; 2000US-0196046.

PA (WHEED) WHITEHEAD INST BIOMEDICAL RES.

PA (ELLI-) ELLIPSIS BIOTHERAPEUTICS CORP.

PI Daly M, Hudson TJ, Lander ES, Rioux J, Siminovitch K;

DR WPI; 2001-367874/38.

PT Testing for the presence of polymorphisms associated with inflammatory
 PT bowel disease, using a hybridization assay -

PS Disclosure; Page 112; 463pp; English.

XX The present invention describes a method for detecting the presence of
 CC polymorphisms associated with inflammatory bowel diseases such as
 CC ulcerative colitis and Crohn's disease. The methods can be used to detect
 CC the presence of genetic polymorphisms associated with inflammatory bowel
 CC disease and correlating their occurrence with disease states. They may be
 CC used in this way for phenotypic correlations, forensics, paternity
 CC testing, medicine and genetic analysis. The present sequence is a gene
 CC containing a polymorphic site described in the exemplification of the
 CC invention.

SQ Sequence 700 BP; 210 A; 141 C; 150 G; 199 T; 0 other;

Query Match

5.2%; Score 52; DB 22; Length 700;
 Best Local Similarity 100.0%; Pred. No. 1.2e-14;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 379 GCCTGTAGTCCAGCTACTTGGAGGCTGAGCGGGAAGAAATCGCTTGAACCC 430
 ||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 589 GCCTGTAGTCCAGCTACTTGGAGGCTGAGCGGGAAGAAATCGCTTGAACCC 640

RESULT 18

AAH92081
 ID AAH92081 standard; DNA; 700 BP.

XX
 AC AAH92081;

DT 09-OCT-2001 (first entry)

DE Human inflammatory bowel disease related gene fragment IGR2093a.

KN Human; inflammatory bowel disease; Crohn's disease; ulcerative colitis;
 KW single nucleotide polymorphism; SNP; chromosome 19p13; paternity test;
 KW chromosome 5q31-33; forensic test; gene therapy; ds.

XX Homo sapiens.

PN WO200142511-A2.

PD 14-JUN-2001.

PF 11-DEC-2000; 2000WO-US33632.

PR 10-DEC-1999; 99US-0170257.

PR 10-APR-2000; 2000US-0196046.

PA (WHEED) WHITEHEAD INST BIOMEDICAL RES.

PA (ELLI-) ELLIPSIS BIOTHERAPEUTICS CORP.

PI Daly M, Hudson TJ, Lander ES, Rioux J, Siminovitch K;

DR WPI; 2001-367874/38.

PT Testing for the presence of polymorphisms associated with inflammatory
 PT bowel disease, using a hybridization assay -

PS Disclosure; Page 112-113; 463pp; English.

XX The present invention describes a method for detecting the presence of
 CC polymorphisms associated with inflammatory bowel diseases such as
 CC ulcerative colitis and Crohn's disease. The methods can be used to detect
 CC the presence of genetic polymorphisms associated with inflammatory bowel
 CC disease and correlating their occurrence with disease states. They may be
 CC used in this way for phenotypic correlations, forensics, paternity
 CC testing, medicine and genetic analysis. The present sequence is a gene
 CC containing a polymorphic site described in the exemplification of the
 CC invention.

SQ Sequence 700 BP; 178 A; 183 C; 188 G; 151 T; 0 other;

Query Match

5.2%; Score 52; DB 22; Length 700;
 Best Local Similarity 100.0%; Pred. No. 1.2e-14;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 379 GCCTGTAGTCCAGCTACTTGGAGGCTGAGCGGGAAGAAATCGCTTGAACCC 430
 ||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 90 GCCTGTAGTCCAGCTACTTGGAGGCTGAGCGGGAAGAAATCGCTTGAACCC 141

RESULT 19

AAL37013/c

ID AAL37013 standard; DNA; 2701 BP.

XX

AC AAL37013;

XX

DT 08-JAN-2002 (first entry)
XX Human musculoskeletal system related polynucleotide SEQ ID NO 3378.
DE Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.
XX
OS Homo sapiens.
XX
XX WO200155367-A1.
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01338.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214986.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234224.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0245474.
PR 08-NOV-2000; 2000US-0245475.
PR 08-NOV-2000; 2000US-0245476.
PR 08-NOV-2000; 2000US-0245477.
PR 08-NOV-2000; 2000US-0245478.
PR 08-NOV-2000; 2000US-0245523.
PR 08-NOV-2000; 2000US-0245524.
PR 08-NOV-2000; 2000US-0245525.
PR 08-NOV-2000; 2000US-0245526.
PR 08-NOV-2000; 2000US-0245527.
PR 08-NOV-2000; 2000US-0245528.
PR 08-NOV-2000; 2000US-0245532.
PR 08-NOV-2000; 2000US-0245609.
PR 08-NOV-2000; 2000US-0245610.
PR 08-NOV-2000; 2000US-0245611.
PR 08-NOV-2000; 2000US-0245613.
PR 17-NOV-2000; 2000US-02459207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.

PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 XX WPT; 2001-451937/48.
 XX
 PT Isolated polypeptide for treating, preventing and/or prognosing
 PT disorders related to the musculoskeletal system including
 PT musculoskeletal cancers and also for testing and detection e.g.
 PT diagnosis -
 XX
 XX Example 2; SEQ ID NO 3378; 781pp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (AAL34669-AAL37666) and proteins
 CC (ABB03087-ABB04109) associated with the musculoskeletal system useful
 CC for preventing, treating or ameliorating medical conditions e.g. by
 CC protein or gene therapy. The genes are isolated from a range of human
 CC tissues disclosed in the specification. The nucleic acids, proteins,
 CC antibodies and (ant)agonists are useful in the diagnosis, treatment
 CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
 CC other cancers of the adrenal gland, bone, bone marrow, breast,
 CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
 CC (c) cardiovascular disorders such as myocardial ischaemias; (d) wound
 CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
 CC and (f) infectious diseases such as viral, bacterial, fungal and
 CC parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 2701 BP; 774 A; 509 C; 580 G; 838 T; 0 other;
 Query Match 5.2%; Score 52; DB 22; Length 2701;
 Best Local Similarity 100.0%; Pred. No. 1.1e-14;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 509 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATTC 560
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||
 942 AAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATTC 891
 RESULT 20
 AAK74315
 ID AAK74315 standard; DNA; 5533 BP.
 XX
 AC AAK74315;
 XX
 DT 07-NOV-2001 (first entry)
 XX
 DE Human Immune/haematopoietic antigen genomic sequence SEQ ID NO:29127.
 XX
 KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis; ds.
 XX
 OS Homo sapiens.
 XX
 XX WC2001517182-A2.
 PN
 XX
 PD 09-AUG-2001.
 XX
 XX

PF 17-JAN-2001; 2001WO-US01354.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225271.
 PR 14-AUG-2000; 2000US-0225275.
 PR 14-AUG-2000; 2000US-0225758.
 PR 18-AUG-2000; 2000US-0225759.
 PR 22-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226688.
 PR 23-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 21-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.

PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 03-JAN-2001; 2001US-0254097.
 PR 03-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 PR Rosen CA, Barash SC, Ruben SM;
 PI WPI: 2001-483426/52.
 DR Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PR

PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Disclosure; SEQ ID NO 29127; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 5533 BP; 1261 A; 1349 C; 1475 G; 1448 T; 0 other;
 Query Match 5.2%; Score 52; DB 22; Length 5533;
 Best Local Similarity 100.0%; Pred. No. 1.1e-14;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 336 ICTCTACTAAATAACAAAATAGCCAGGTGGTGGCACACGCCCTGTAGT 387
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 3981 TCTCTACTAAATAACAAAATAGCCAGGTGGTGGCACACGCCCTGTAGT 4032
 RESULT 21
 AAK74317
 ID AAK74317 standard; DNA; 5533 BP.
 AC AAK74317;
 XX
 XX 07-NOV-2001 (first entry)
 XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:29129.
 XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 XX Cytostatic; gene therapy; vaccine; metastasis; ds.
 OS Homo sapiens.
 XX
 XX WO200157182-A2.
 PN
 PD 09-AUG-2001.
 XX
 XX 17-JAN-2001; 2001WO-US01354.
 XX
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 14-JUL-2000; 2000US-0217496.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234597.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.

PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis -

Disclosure; SEQ ID NO 29129; 3071pp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I) amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patient's own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/hematopoietic-related diseases, especially cancers and cancer metastases of hematopoietic-derived cells. AAK64703 to AAK87694 represent human immune/hematopoietic antigen genomic

CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention.

SQ Sequence 5533 BP; 1361 A; 1349 C; 1475 G; 1448 T; 0 other;

Query Match 5.28; Score 52; DB 22; Length 5533;
Best Local Similarity 100.0%; Pred. No. 1.le-14;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 336 TCTCTCTAAATAACAAAAATTAGCCAGGTGTGGTGCCACGCCCTGTACT 387
|||||

Db 3981 TCTCTCTAAATAACAAAAATTAGCCAGGTGTGGTGCCACGCCCTGTACT 4032

RESULT 22

AAK74316

ID AAK74316 standard; DNA; 5536 BP.

XX

AC AAK74316;

XX

DT 07-NOV-2001 (first entry)

XX

DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:29128.

XX

KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

XX

KW cytostatic; gene therapy; vaccine; metastasis; ds.

XX

OS Homo sapiens.

XX

PN WO200157182-A2.

XX

PD 09-AUG-2001.

XX

PF 17-JAN-2001; 2001WO-US01354.

XX

PR 31-JAN-2000; 2000US-0179065.

PR

PR 04-FEB-2000; 2000US-0180628.

PR

PR 24-FEB-2000; 2000US-0184664.

PR

PR 02-MAR-2000; 2000US-0186350.

PR

PR 16-MAR-2000; 2000US-0189874.

PR

PR 17-MAR-2000; 2000US-0190076.

PR

PR 18-APR-2000; 2000US-0198123.

PR

PR 19-MAY-2000; 2000US-0205515.

PR

PR 07-JUN-2000; 2000US-0209467.

PR

PR 28-JUN-2000; 2000US-0214886.

PR

PR 30-JUN-2000; 2000US-0215135.

PR

PR 07-JUL-2000; 2000US-0216647.

PR

PR 11-JUL-2000; 2000US-0216880.

PR

PR 11-JUL-2000; 2000US-0217487.

PR

PR 14-JUL-2000; 2000US-0217496.

PR

PR 26-JUL-2000; 2000US-0218290.

PR

PR 26-JUL-2000; 2000US-0220963.

PR

PR 14-AUG-2000; 2000US-0224518.

PR

PR 14-AUG-2000; 2000US-0224519.

PR

PR 14-AUG-2000; 2000US-0225213.

PR

PR 14-AUG-2000; 2000US-0225214.

PR

PR 14-AUG-2000; 2000US-0225266.

PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249309.
PR 17-NOV-2000; 2000US-0249310.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.

PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
XX
PS Disclosure; SEQ ID NO 29128; 3071pp + Sequence Listing; English.
XX
PS
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytosolic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 5536 BP; 1263 A; 1348 C; 1474 G; 1451 T; 0 other;

Query Match 5.2%; Score 52; DB 22; Length 5536;
Best Local Similarity 100.0%; Pred. No. 1.1e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 336 TCCTACTAAATAACAAATAGCCAGGTGTGGTGGCACACGCCCTGTAGT 387
|||||
DB 3984 TCCTACTAAATAACAAATAGCCAGGTGTGGTGGCACACGCCCTGTAGT 4035
|||||

RESULT 23
AAK74318
ID AAK74318 standard; DNA; 5537 BP.
XX
AC AAK74318;
XX

DT 07-NOV-2001 (first entry)
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:29130.
DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
KW Homo sapiens.
XX
XX WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.

```
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-024617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249266.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251866.
PR 08-DEC-2000; 2000US-0251867.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PR XX
PR PA (HUMA-) HUMAN GENOME SCI INC.
PR XX
PR XX Rosen CA, Barash SC, Ruben SM;
PR PI WPI; 2001-483426/52.
PR XX
PR DR Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PR XX useful for preventing, diagnosing and/or treating cancers and
PR PT metastasis -
PR XX
PR XX Disclosure; SEQ ID NO 29130; 3071pp + Sequence Listing; English.
PR PS
PR CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
PR CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
PR CC activity, and can be used in gene therapy and vaccine production. (I)
PR CC proteins and polynucleotides may be used in the prevention, diagnosis and
PR CC treatment of diseases associated with inappropriate (I) expression. For
PR CC example, they may be used to treat disorders associated with decreased
PR CC expression by rectifying mutations or deletions in a patient's genome
PR CC that affect the activity of (I) by expressing inactive proteins or to
PR CC supplement the patient's own production of (I). Additionally, (I)
PR CC polynucleotides may be used to produce the secreted (I), by inserting
PR CC the nucleic acids into a host cell and culturing the cell to express the
PR CC protein. (I) proteins and polynucleotides may be used to prevent,
PR CC diagnose and treat immune/hematopoietic-related diseases, especially
PR CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
PR CC to AAK37694 represent human immune/hematopoietic antigen genomic
PR CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
PR CC represent sequences used in the exemplification of the present invention.
PR XX
PR SQ Sequence 5537 BP; 1271 A; 1331 C; 1454 G; 1481 T; 0 other;
Query Match 5.2%; Score 52; DB 22; Length 5537;
Best Local Similarity 100.0%; Pred. No. 1.1e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 336 TCTCTACTAAAAATACAAAATTAGCCAGGTGTGGTGACACACGCTGTAGT 387
Db 3988 TCTCTACTAAAAATACAAAATTAGCCAGGTGTGGTGACACACGCTGTAGT 4039
RESULT 24
AAK89991/c
ID AAK89991 standard; DNA; 12244 BP.
XX
XX AAK89991;
AC
XX
XX 05-NOV-2001 (first entry)
XX
XX Human digestive system antigen genomic sequence SEQ ID NO: 3567.
XX
XX Human; digestive system antigen; gene therapy; cancer; appendicitis;
XX ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
XX digestive system disorder; Meckel's diverticulum; ds.
XX
XX Homo sapiens.
XX
XX WO200155314-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01324.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
```

PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0190076.
 PR 19-MAY-2000; 2000US-0198123.
 PR 07-JUN-2000; 2000US-0205515.
 PR 28-JUN-2000; 2000US-0209467.
 PR 30-JUN-2000; 2000US-0214886.
 PR 07-JUL-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 11-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 14-AUG-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 18-AUG-2000; 2000US-0225759.
 PR 22-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 01-SEP-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0233968.
 PR 14-SEP-2000; 2000US-0233977.
 PR 14-SEP-2000; 2000US-0233998.
 PR 14-SEP-2000; 2000US-0233999.
 PR 14-SEP-2000; 2000US-0234000.
 PR 14-SEP-2000; 2000US-0234201.
 PR 14-SEP-2000; 2000US-0234063.
 PR 14-SEP-2000; 2000US-0234064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 25-SEP-2000; 2000US-0234999.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249300.
 PR 17-NOV-2000; 2000US-0249309.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-502630/55.

Polynucleotides encoding digestive system antigens, useful for diagnosing, treating, preventing and/or prognosing disorders of the digestive system, particularly cancer and cancer metastases -

Disclosure; SEQ ID NO 3567; 986pp; English.

The present invention provides the protein and coding sequences of a

CC number of human digestive system antigens. These can be used in the
 CC diagnosis, treatment and prevention of digestive system disorders,
 CC including cancer, Meckel's diverticulum, bacterial or parasitic
 CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
 CC ulcerative colitis. The present sequence is a genomic DNA fragment
 CC encoding a digestive system antigen of the invention.
 XX
 SQ Sequence 12244 BP; 3636 A; 2432 C; 2513 G; 3663 T; 0 other;

Query Match 5.2%; Score 52; DB 22; Length 12244;
 Best Local Similarity 100.0%; Pred. No. 1.le-14;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 GATCAGTTCAGGTCAGGAGTTTCAGACGAGCCTGGCCACACGGTGAAGC 197
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 10382 GATCAGTTCAGGTCAGGAGTTTCAGACGAGCCTGGCCACACGGTGAAGC 10311

RESULT 25

ABAL5896/C

ID ABAL5896 standard; DNA, 12392 BP.

XX AC ABAL5896;

XX DT 23-JAN-2002 (first entry)

XX DE Human nervous system related polynucleotide SEQ ID NO 8227.

XX KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
 KW antiparkinsonian; antisking; antianemic; antiarthritic; cancer;
 KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
 KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
 KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.

XX OS Homo sapiens.

XX PN WC200159063-A2.

XX PD 16-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01334.
 XX PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0203515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 14-JUL-2000; 2000US-0217496.
 PR 26-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 14-AUG-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 23-AUG-2000; 2000US-0227182.
 PR 30-AUG-2000; 2000US-0227009.
 PR 01-SEP-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 20-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 20-OCT-2000; 2000US-0242221.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.

PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-541565/60.
XX
PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system
PT cancers and metastases -
XX
PS Disclosure; SEQ ID NO 8227; 1701pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (ABBI4678-ABBI8001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 12392 BP; 3812 A; 2126 C; 2413 G; 4041 T; 0 other;
Query Match 5.2%; Score 52; DB 22; Length 12392;
Best Local Similarity 100.0%; Pred. No. 1.1e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 174 CAGCCTGGCCACACGGTGAAACCCCATCTCTACTAAAAATAAAAAATTAGC 225

Db 3234 CAGCCTGGCCACACGGTGAAACCCCATCTCTACTAAAAATAAAAAATTAGC 3183
RESULT 26
AAK95240/C
ID AAK95240 standard; DNA; 1503900 BP.
XX
AC AAK95240;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human neuregulin-1 gene.
XX
KW Human; neuregulin-1 associated gene 1; NRGIAG1; Schizophrenia gene;
KW gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200164876-A2.
XX
PD 07-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US06376.
XX
PR 28-FEB-2000; 2000US-0515715.
XX
PR (DECO-) DECODE GENETICS EHF.
XX
PI Stefansson H, Steinthorsdottir V, Gulcher JR;
XX
XX WPI; 2001-550179/61.
DR P-PSDB; AAG67900, AAG67901, AAG67902, AAG67903, AAG67904, AAG67905,
DR AAG67906, AAG67907, AAG67908, AAG67909, AAG67910, AAG67911, AAG67912,
DR AAG67913, AAG67914, AAG67915, AAG67916, AAG67917, AAG67918, AAG67919,
DR AAG67920, AAG67921, AAG67922, AAG67923, AAG67924, AAG67925, AAG67926,
DR AAG67927, AAG67928, AAG67929, AAG67930, AAG67931, AAG67932, AAG67933,
DR AAG67934, AAG67935, AAG67936, AAG67937.
XX
PT Neuregulin-1 associated gene 1 nucleic acids and fragments, useful for
PT preventing diagnosing and treating schizophrenia -
XX
PS Disclosure; Page 90-501; 750pp; English.
XX
CC This sequence represents the human neuregulin-1 associated gene 1
CC (NRGIAG1) of the invention. The NRGIAG1 gene is also referred to as the
CC human Schizophrenia gene. The invention also relates to fragments or
CC variants of the gene and the NRGIAG1 polypeptides they encode. The
CC NRGIAG1 nucleic acids and polypeptides may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate NRGIAG1
CC expression. For example, they may be used to treat disorders associated
CC with decreased expression by rectifying mutations or deletions in a
CC patient's genome that affect the activity of NRGIAG1 by expressing
CC inactive proteins or to supplement the patients own production of
CC NRGIAG1. Additionally, the gene may be used to produce NRGIAG1
CC polypeptides, by inserting the nucleic acids into a host cell and
CC culturing the cell to express the protein. The gene may also be used as
CC DNA probes and primers in diagnostic assays to detect and quantitate the
CC presence of similar nucleic acids in samples, and therefore which
CC patients may be in need of restorative therapy. The NRGIAG1 polypeptides
CC may also be used as antigens in the production of antibodies against
CC NRGIAG1 and in assays to identify modulators of NRGIAG1 expression and
CC activity. Anti-NRGIAG1 antibodies and antagonists may also be used to
CC down regulate expression and activity. Anti-NRGIAG1 antibodies may
CC also be used as diagnostic agents for detecting the presence of NRGIAG1
CC polypeptides in samples. NRGIAG1 is associated with schizophrenia which
CC may be prevented, diagnosed and/or treated by the above methods.
XX
SQ Sequence 1503900 BP; 452487 A; 281874 C; 288074 G; 480092 T; 1373 other;
Query Match 5.2%; Score 52; DB 22; Length 1503900;
Best Local Similarity 100.0%; Pred. No. 9.4e-15;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible]

RESULT 27
AAK96733/C
ID AAK96733 standard; DNA; 1503900 BP.
XX
AC AAK96733:

DT	17-DEC-2001	(first entry)
XX	Human	
DE	Human	neuregulin-1 gene.
XX		
KW	Human;	neuregulin 1 gene; schizophrenia; gene therapy; ds.
XX		
OS	Homo sapiens.	
XX		
PN	W0200164877-A2.	
XX		
PD	07-SEP-2001.	

28-FEB-2001; 2001WO-US06377.
 PF XX
 28-FEB-2000; 2000US-0515716.
 PR XX
 28-FEB-2000; 2000US-0515716.
 PR XX
 (DECO-) DECODE GENETICS EHF.
 PA

Stefansson H, Steinthorsdottir V, Gulcher JR;
WPI: 2001-514841/56.
P-PSDB: DR DR DR DR DR DR
RAG67938, RAG67939, RAG67940, RAG67941, RAG67942,
RAG67944, RAG67945, RAG67946, RAG67947, RAG67948,
RAG67950, RAG67951, RAG67952, RAG67953, RAG67949
RAG67956, RAG67957, RAG67958, RAG67959
RAG67962, RAG67963, RAG67964, RAG67965,
RAG67968, RAG67969, RAG67970, RAG67971, RAG67972,
RAG67974, RAG67975

PT Neuregulin 1 nucleic acids and proteins useful for diagnosing
PT preventing and treating schizophrenia - XX
PS Disclosure; Page 345-756; 756pp; English.

This sequence represents the human neurotrophin 1 gene of the invention. The invention also relates to fragments or variants of the neurotrophin 1 gene. The gene and its proteins may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate neurotrophin 1 expression, such as schizophrenia. For example they may be used to treat disorders associated with decreased neurotrophin 1 expression by rectifying mutations or deletions in a patient's genome that affect the activity of neurotrophin 1 by expressing inactive proteins or to supplement the patients own production of polypeptides. Additionally, the gene may be used to produce the neurotrophin 1 protein, by inserting the nucleic acids into a host cell and culturing the cell to express the protein. The gene and its complementary sequences may also be used as DNA probes in diagnostic assays to detect and quantify the presence of similar nucleic acids in samples and therefore which patients may be in need of restorative therapy. The protein may also be used as antigens in the production of antibodies against neurotrophin 1 and in assays to identify modulators of neurotrophin 1 expression and activity. The antibodies and antagonists may also be used to down regulate expression and activity. The antibodies may also be used as diagnostic agents for detecting the presence of neurotrophin 1 in samples.

Query Match 5.2%; Score 52; DB 22; Length 1503900;
 Best Local Similarity 100.0%; Pred. No. 9.4e-15;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps

QY 509 AAAATAAAATAAAAATAAAATAAAATAAAATAAAATAAAATTC 560
|||||
Db 1366726 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATTC 1366675

RESULT 28
ABN80120
ID ABN80120 standard; DNA; 5523 BP.
XX
XX AC ABN80120;
XX AC ABN80120;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human chemically modified disease associated gene SEQ ID NO 137.

KW Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;
 KW heart disease; epilepsy; histone deacetylation; muscular dystrophy;
 KW dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;
 KW antidiabetic; cytostatic; anticonvulsant; ds.

OS	Homo sapiens.
OS	Synthetic

PN WO200200927-A2.

PD 03-JAN-2002.

02-III-2001.

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

PR 01-SEP-2000; 2000DE-1043826.

PA (EPIG-) EPIGENOMICS AG.

PT 01ek 3 Dienenbrook c

[illegible]

Novel nucleic acid useful for diagnosis and therapy of diseases associated with development genes such as diabetes, comprises a sequence of a segment of chemically pretreated DNA of genes associated with development.

PS Claim 1; SEQ ID NO 137; 27pp; English.

The invention relates to a nucleic acid (I) comprising a sequence at least 13 bases in length of a segment of chemically pretreated DNA (II) of genes associated with development selected from 87 genes listed in the specification such as ACCPN, ADPN, or AFD1 and comprising one of 350 sequences (ABN79984-ABN80333) or their complements. The invention is useful for the diagnosis or therapy of diseases associated with development genes, in particular disease related to homeobox containing genes (HOX), like diabetes, cancer, apoptosis related diseases, syndromes associated with congenital heart disease, epilepsy, diseases related to histone deacetylation, Curranio syndrome, diseases related with the development of the brain and limb girdle muscular dystrophy and dwarfism. Oligomers specific to each of the genes are useful for detecting the methylation state of all CpG dinucleotides within the 350 sequences or (II) and their complementary sequences, as primer oligonucleotides for the amplification of the 350 sequences, (II) and/or their complements and as oligomer probes for detecting the cytosine methylation state and/or single nucleotide polymorphisms (SNPs).
Note: The sequence data for this patent did not form part of the printed specification but is based on sequence information supplied to Derwent by the European Patent Office.

Sequence 6523 BP; 1499 A; 356 C; 1860 G; 2808 T; 0 other;

Query Match 5.1%; Score 51; DB 24; Length 6523;
Best Local Similarity 100.0%; Pred. No. 3.2e-14;
Matches 51; Conservative 0.0; Mismatches 0.0; T-31

509 AAAATATAAATAAAATATAAATAAAATATAAATAAAATTT 559

```
Db 3112 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATT 3162
|||||
RESULT 29
ABQ88169
ID ABQ88169 standard; cDNA; 43411 BP.
XX AC ABQ88169;
XX DT 18-SEP-2002 (first entry)
XX DE Human osteoblast differentiation related cDNA SEQ ID NO 76.
XX KW Human; osteoblast; stem cell differentiation; bone tissue deposition;
XX KW osteoporosis; osteopathic; ss.
XX OS Homo sapiens.
XX PN WO200250301-A2.
XX PD 27-JUN-2002.
XX PF 18-DEC-2001; 2001WO-US48276.
XX PR 18-DEC-2000; 2000US-255882P.
XX PR 24-APR-2001; 2001US-285691P.
XX PA (GENE-) GENE LOGIC INC.
XX PA (PROC ) PROCTER & GAMBLE CO.
XX PI Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;
XX PI Mertz L;
XX DR WPI; 2002-557663/59.
XX
XX Use of genes and their expression profiles associated with osteoblast
PT differentiation for screening modulators bone formation, for diagnosing
PT or treating e.g. osteoporosis, or as markers for the differentiation
PT process
XX
PS Claim 1; SEQ ID NO 76; 78pp + Sequence Listing; English.
XX
CC The invention relates to genes and their expression profiles are used
CC for:
CC (a) screening modulators of precursor stem cell differentiation into
CC osteoblasts, or bone tissue deposition;
CC (b) diagnosing abnormal deposition of bone tissue, abnormal rate of
CC osteoblast formation or osteoporosis; or
CC (c) treating or monitoring treatment of the conditions cited in (b), or
CC monitoring the progression of bone tissue deposition.
CC Specific conditions include postmenopausal osteoporosis, glucocorticoid
CC osteoporosis or male osteoporosis, osteopenia, osteodystrophy,
CC that involve altered bone metabolism (e.g. idiopathic juvenile
CC Fanconi syndrome or fibrous dysplasia. The present sequence is that of an
CC osteoblast differentiation associated cDNA marker of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 43411 BP; 9410 A; 12601 C; 12923 G; 8477 T; 0 other;
Query Match 5.1%; Score 51; DB 24; Length 43411;
Best Local Similarity 100.0%; Pred. No. 3.1e-14;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 335 GTCTTACTAAATAACAAATAATAGCCAGGTGTGTGTCACACGCGCTGTA 385
|||||
DB 944 GTCTTACTAAATAACAAATAATAGCCAGGTGTGTGTCACACGCGCTGTA 994
```

```
RESULT 30
ABK83497/C
ID ABK83497 standard; cDNA; 227968 BP.
XX AC ABK83497;
XX DT 14-AUG-2002 (first entry)
XX DE Human cDNA differentially expressed in granulocytic cells #68.
XX DE ss; granulocytic cell; DNA chip; bacterial infection;
XX KW viral infection; parasitic infection; protozoal infection;
XX KW fungal infection; sterile inflammatory disease; psoriasis;
XX KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
XX KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
XX KW adult respiratory distress syndrome; inflammatory bowel disease;
XX KW Crohn's disease; ulcerative colitis; periodontal disease;
XX KW granulocyte activation; chronic inflammation; allergy.
XX OS Homo sapiens.
XX PN WO200228999-A2.
XX PD 11-APR-2002.
XX PF 03-OCT-2001; 2001WO-US30821.
XX PR 03-OCT-2000; 2000US-237189P.
XX PA (GENE-) GENE LOGIC INC.
XX PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX PI WPI; 2002-435328/46.
XX
XX Detecting granulocyte activation by detecting differential expression
PT of genes associated with granulocyte activation, which serves as
PT diagnostic markers that is useful for monitoring disease states and
PT drug toxicity
XX
PS Claim 1; SEQ ID NO 68; 114pp; English.
XX
CC The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification and comparing
CC the expression level to an expression level in an unactivated
CC GC, where differential expression of Gs is indicative of GCA.
CC Also included are modulating (M2) GA by contacting GC with an agent
CC that alters the expression of at least one gene in Gs; (2) screening (M3)
CC for an agent capable of modulating GCA or an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease using the
CC gene expression profile; (3) detecting (M4) an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease, by detecting the
CC level of expression in a sample of the tissue of gene(s) from Gs, where
CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
CC an allergic response in a subject, exposure of a subject to a pathogen
CC or sterile inflammatory disease, by contacting a tissue having
CC inflammation with an agent that modulates the expression of gene(s)
CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
CC modulating GCA; M3 is useful for screening an agent capable of modulating
CC GCA preferably in an inflammation in a tissue; M4 is useful for
CC detecting an inflammation (especially chronic) in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
CC reperfusion injury, ARDS, adult respiratory distress syndrome,
CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
CC periodontal disease; also bacterial infection, viral infection,
CC parasitic infection, protozoal infection, fungal infection and M5 is
CC useful for treating one of the above conditions. The present
```

CC sequence represents a gene differentially expressed in granulocytes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pot_sequences.

SQ Sequence 227968 BP; 53185 A; 60440 C; 61985 G; 52358 T; 0 other;

Query Match 5.1%; Score 51; DB 24; Length 227968;
Best Local Similarity 100.0%; Pred. No. 3e-14;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 335 GTCCTCTACTAAAATACAAAATAGCCAGGTGGTGGGCACACGCTGTA 385
|||||
Db 47716 GTCCTCTACTAAAATACAAAATAGCCAGGTGGTGGGCACACGCTGTA 47666

RESULT 31

AAS36586/C

ID AAS36586 standard; DNA; 276 BP.

XX AC AAS36586;

XX XX

DT 17-DEC-2001 (first entry)

XX XX

DE Human cardiovascular system antigen genomic DNA SEQ ID No 2086.

XX XX

KW Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;
KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
KW cerebroprotective; nontropic; antibacterial; virucide; fungicide; cancer;
KW ophthalmological; vulnary; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; tissue regeneration;
KW anti-infertility.

OS Homo sapiens.

XX XX

PN WO200153321-A2.

XX XX

PD 02-AUG-2001.

XX XX

PF 17-JAN-2001; 2001WO-US01340.

XX XX

PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184564.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 11-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 14-AUG-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226581.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 03-SEP-2000; 2000US-0229509.
PR 03-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.

PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-451930/48.
DR
XX
XX New cardiovascular system related polynucleotides and polypeptides,
PT useful for diagnosing, treating and/or preventing disorders of the
PT cardiovascular system -
PT
XX
PS Claim 1; SEQ ID No 2086; 674pp; English.
XX
CC Sequences AAS35741-AAS36942 represent genomic DNA molecules, which encode
CC the cardiovascular system antigens and their associated polynucleotides are
CC useful in the diagnosis, treatment and prevention of various types of
CC disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
CC chickens or sheep. A pathological condition can be determined by
CC detecting the presence or absence of a mutation in a cardiovascular
CC system antigen polynucleotide. The treatable disorders include autoimmune
CC diseases such as rheumatoid arthritis, hyperproliferative disorders such
CC as neoplasms of the breast or liver, cardiovascular disorders such as
CC cardiac arrest, cerebrovascular disorders such as cerebral ischaemia,
CC nervous system disorders such as Alzheimer's disease, infections caused
CC by bacteria, viruses and fungi, ocular disorders such as corneal
CC infection, endocrine disorders such as premature labour and infertility,
CC gastrointestinal disorders such as Crohn's disease, renal disorders such
CC as glomerulonephritis and respiratory disorders such as asthma and
CC pleurisy. The polypeptides can also be used to aid wound healing, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, to regenerate tissues and in chemotaxis.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

Query Match 5.0%; Score 50; DB 22; Length 276;
Best Local Similarity 100.0%; Pred. No. 1.1e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 166 TTTGAGACCGCTGGCCACACGGTGAACCCCATCTCTACTAAATA 215
Db 216 TTTGAGACCGCTGGCCACACGGTGAACCCCATCTCTACTAAATA 167

RESULT 32
AAK75295/c
ID AAK75295 standard; DNA; 276 BP.
XX
AC AAK75295;
XX
XX 07-NOV-2001 (first entry)
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:30107.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW Cytostatic; gene therapy; vaccine; metastasis; ds.
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.

KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
XX WO200157182-A2.
PN 09-AUG-2001.
PD 17-JAN-2001; 2001WO-US01354.
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225366.
PR 14-AUG-2000; 2000US-0225367.
PR 14-AUG-2000; 2000US-0225467.
PR 14-AUG-2000; 2000US-0225468.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0244674.
PR 08-NOV-2000; 2000US-0245475.
PR 08-NOV-2000; 2000US-0245476.
PR 08-NOV-2000; 2000US-0245477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.

```
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX
XX Disclosure; SEQ ID NO 30108; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAK62170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patient's own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/haematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/haematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention.
XX
XX Sequence 276 BP; 50 A; 81 C; 69 G; 76 T; 0 other;
SQ
Query Match 5.0%; Score 50; DB 22; Length 276;
Best Local Similarity 100.0%; Pred. No. 1.1e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 166 TTGTGAGCAGCGTGGCCACACGGTGAACCCGATCTCTACTAAATA 215
DB 216 TTGTGAGCAGCGTGGCCACACGGTGAACCCGATCTCTACTAAATA 167
RESULT 34
ABV54360
ID ABV54360 standard; cDNA; 475 BP.
XX
XX AC ABV54360;
XX
XX DT 17-SEP-2002 (first entry)
XX
XX DE Human prostate expression marker cDNA 54351.
XX
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX OS Homo sapiens.
XX
XX PN W0200160860-A2.
XX
XX PD 23-AUG-2001.
XX
XX PF 20-FEB-2001; 2001WO-US05171.
XX
XX PR 17-FEB-2000; 2000US-183319P.
XX PR 16-MAR-2000; 2000US-189862P.
XX PR 25-MAY-2000; 2000US-207454P.
XX PR 09-JUN-2000; 2000US-211314P.
XX PR 18-JUL-2000; 2000US-219007P.
XX PR 13-DEC-2000; 2000US-255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
```

```
PI Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer -
XX
XX Claim 1; Page 10509; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for:
XX (a) assessing whether a patient is afflicted with prostate cancer;
XX (b) monitoring the progression of prostate cancer in a patient;
XX (c) assessing the efficacy of a test compound to inhibit prostate
XX cancer in a patient;
XX (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound;
XX (g) determining whether prostate cancer has metastasized in a patient;
XX (h) assessing the aggressiveness or indolence of prostate cancer in a
XX patient;
XX (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 475 BP; 180 A; 62 C; 80 G; 133 T; 0 other;
SQ
Query Match 5.0%; Score 50; DB 23; Length 475;
Best Local Similarity 100.0%; Pred. No. 1e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 509 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 558
DB 193 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 242
RESULT 35
ABV55359/c
ID ABV55359 standard; cDNA; 578 BP.
XX
XX AC ABV55359;
XX
XX DT 17-SEP-2002 (first entry)
XX
XX DE Human prostate expression marker cDNA 55350.
XX
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX OS Homo sapiens.
XX
XX PN W0200160860-A2.
XX
XX PD 23-AUG-2001.
XX
XX PF 20-FEB-2001; 2001WO-US05171.
XX
XX PR 17-FEB-2000; 2000US-183319P.
XX PR 16-MAR-2000; 2000US-189862P.
XX PR 25-MAY-2000; 2000US-207454P.
XX PR 09-JUN-2000; 2000US-211314P.
XX PR 18-JUL-2000; 2000US-219007P.
XX PR 13-DEC-2000; 2000US-255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX PI Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
```

PT for detecting presence of prostate cancer, stage of prostate cancer -
XX Claim 1; Page 10692-10693; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 578 BP; 160 A; 121 C; 85 G; 211 T; 1 other;

Query Match 5.0%; Score 50; DB 23; Length 578;
Best Local Similarity 100.0%; Pred. No. 1e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 509 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 558
|||||
DB 379 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 330

RESULT 36
AAI94916
ID AAI94916 standard; CDNA; 737 BP.
XX
AC AAI94916;
XX
DT 13-NOV-2001 (first entry)
XX
DE Human neuroblastoma expressed polynucleotide SEQ ID NO 991.
XX
KW Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.
XX
OS Homo sapiens.
XX
PN W0200166719-A1.
XX
PD 13-SEP-2001.
XX
PF 02-MAR-2001; 2001WO-JP01629.
XX
PP 07-MAR-2000; 2000JP-0159195.
XX
PA (CHIB-) CHIBA PREFECTURE.
PA (HISM-) HISAMITSU PHARM CO LTD.
XX
PI Nakagawara A;
XX
PI WPI; 2001-565584/63.
XX
DR Nucleic acids originating in gene expressed in human neuroblastoma,
PT useful as probe or primer in diagnosing prognosis of human
PT neuroblastoma, malignancy and susceptibility indicator or tumour marker
PT for anti-cancer agents -
XX
PS Claim 1; Page 763-764; 2979pp; Japanese.
XX
CC The invention relates to novel genes (AAI93926-AAI97963) expressed in
CC human neuroblastoma. The nucleic acids are applicable as a probe or
CC primer in diagnosing the prognosis of human neuroblastoma, malignancy and
CC susceptibility indicators or tumour markers for anti-cancer agents. The
CC gene information for diagnosing prognosis is related to factors similar
CC to that for N-myc and TrkA genes.

XX
SQ Sequence 737 BP; 218 A; 140 C; 140 G; 216 T; 23 other;

Query Match 5.0%; Score 50; DB 22; Length 737;
Best Local Similarity 100.0%; Pred. No. 1e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 509 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 558
|||||
DB 361 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 410

RESULT 37
ABA16762/c
ID ABA16762 standard; DNA; 4243 BP.
XX
AC ABA16762;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 9093.
XX
DE
XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antischlicking; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antilucer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN W0200159063-A2.
XX
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01334.
XX
PP 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227182.

PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0233055.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 20-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0242221.
PR 08-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.

PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 03-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
XX useful for preventing, diagnosing and/or treating nervous system
XX cancers and metastases -
XX
XX Disclosure; SEQ ID NO 9093; 1701pp + Sequence Listing; English.

XX The invention relates to novel genes (AB11004-ABA21534) and proteins
XX (AB114678-AB118001) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful
XX in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
XX and ovarian cancer and other cancers of the adrenal gland, bone, bone
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
XX (b) immune disorders e.g. Addison's disease, allergies, autoimmune
XX haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative
XX colitis; (c) cardiovascular disorders such as myocardial ischaemia;
XX (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
XX and parasitic infections.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 4243 BP; 1117 A; 967 C; 1007 G; 1152 T; 0 other;

Query Match 5.0%; Score 50; DB 22; Length 4243;
Best Local Similarity 100.0%; Pred. No. 9.8e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 TTTCAGACAGCTGGCCACACAGCGTGAACCCATCTCTACTAAATA 215
|||||
Db 1583 TTTCAGACAGCTGGCCACACAGCGTGAACCCATCTCTACTAAATA 1534

RESULT 38
AAK74563

ID XX AAK74563 standard; DNA; 5283 BP.
AC XX AAK74563;
XX XX
DT DT 07-NOV-2001 (first entry)
XX XX
XX XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:29375.
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW KW Cytostatic; gene therapy; vaccine; metastasis; ds.
XX XX
OS OS Homo sapiens.
XX XX
XX XX WO200157182-A2.
XX XX
XX XX 09-AUG-2001.
XX XX
XX XX 17-JAN-2001; 2001WO-US01354.
XX XX
XX XX 31-JAN-2000; 2000US-0179065.
XX XX 04-FEB-2000; 2000US-0180628.
XX XX 24-FEB-2000; 2000US-0184664.
XX XX 02-MAR-2000; 2000US-0186350.
XX XX 16-MAR-2000; 2000US-0189874.
XX XX 17-MAR-2000; 2000US-0190076.
XX XX 18-APR-2000; 2000US-0198123.
XX XX 19-MAY-2000; 2000US-0205515.
XX XX 07-JUN-2000; 2000US-0209467.
XX XX 28-JUN-2000; 2000US-0214886.
XX XX 30-JUN-2000; 2000US-0215135.
XX XX 07-JUL-2000; 2000US-0216647.
XX XX 07-JUL-2000; 2000US-0216880.
XX XX 11-JUL-2000; 2000US-0217487.
XX XX 11-JUL-2000; 2000US-0217496.
XX XX 14-JUL-2000; 2000US-0218290.
XX XX 26-JUL-2000; 2000US-0220963.
XX XX 26-JUL-2000; 2000US-0220964.
XX XX 14-AUG-2000; 2000US-0224518.
XX XX 14-AUG-2000; 2000US-0224519.
XX XX 14-AUG-2000; 2000US-0225213.
XX XX 14-AUG-2000; 2000US-0225214.
XX XX 14-AUG-2000; 2000US-0225266.
XX XX 14-AUG-2000; 2000US-0225267.
XX XX 14-AUG-2000; 2000US-0225268.
XX XX 14-AUG-2000; 2000US-0225270.
XX XX 14-AUG-2000; 2000US-0225447.
XX XX 14-AUG-2000; 2000US-0225757.
XX XX 14-AUG-2000; 2000US-0225758.
XX XX 14-AUG-2000; 2000US-0225759.
XX XX 18-AUG-2000; 2000US-0226279.
XX XX 22-AUG-2000; 2000US-0226681.
XX XX 22-AUG-2000; 2000US-0226868.
XX XX 23-AUG-2000; 2000US-0227182.
XX XX 23-AUG-2000; 2000US-0227009.
XX XX 30-AUG-2000; 2000US-0228924.
XX XX 01-SEP-2000; 2000US-0229287.
XX XX 01-SEP-2000; 2000US-0229343.
XX XX 01-SEP-2000; 2000US-0229344.
XX XX 01-SEP-2000; 2000US-0229345.
XX XX 05-SEP-2000; 2000US-0229509.
XX XX 05-SEP-2000; 2000US-0229513.
XX XX 06-SEP-2000; 2000US-0230437.
XX XX 08-SEP-2000; 2000US-0231242.
XX XX 08-SEP-2000; 2000US-0231243.
XX XX 08-SEP-2000; 2000US-0231244.
XX XX 08-SEP-2000; 2000US-0231413.
XX XX 08-SEP-2000; 2000US-0231414.
XX XX 08-SEP-2000; 2000US-0232080.
XX XX 08-SEP-2000; 2000US-0232081.
XX XX 12-SEP-2000; 2000US-0231968.
XX XX 14-SEP-2000; 2000US-0232397.
XX XX 14-SEP-2000; 2000US-0232398.
XX XX 14-SEP-2000; 2000US-0232399.
XX XX 14-SEP-2000; 2000US-0232400.
XX XX 14-SEP-2000; 2000US-0232401.
XX XX 14-SEP-2000; 2000US-0233063.
XX XX 14-SEP-2000; 2000US-0233064.
XX XX 14-SEP-2000; 2000US-0233065.
XX XX 21-SEP-2000; 2000US-0234223.
XX XX 21-SEP-2000; 2000US-0234274.
XX XX 25-SEP-2000; 2000US-0234997.
XX XX 25-SEP-2000; 2000US-0234998.
XX XX 25-SEP-2000; 2000US-0235484.
XX XX 27-SEP-2000; 2000US-0235834.
XX XX 27-SEP-2000; 2000US-0235836.
XX XX 29-SEP-2000; 2000US-0236327.
XX XX 29-SEP-2000; 2000US-0236367.
XX XX 29-SEP-2000; 2000US-0236368.
XX XX 29-SEP-2000; 2000US-0236369.
XX XX 29-SEP-2000; 2000US-0236370.
XX XX 02-OCT-2000; 2000US-0236802.
XX XX 02-OCT-2000; 2000US-0237037.
XX XX 02-OCT-2000; 2000US-0237038.
XX XX 02-OCT-2000; 2000US-0237039.
XX XX 02-OCT-2000; 2000US-0237040.
XX XX 13-OCT-2000; 2000US-0239935.
XX XX 13-OCT-2000; 2000US-0239937.
XX XX 20-OCT-2000; 2000US-0240960.
XX XX 20-OCT-2000; 2000US-0241221.
XX XX 20-OCT-2000; 2000US-0241785.
XX XX 20-OCT-2000; 2000US-0241786.
XX XX 20-OCT-2000; 2000US-0241787.
XX XX 20-OCT-2000; 2000US-0241808.
XX XX 20-OCT-2000; 2000US-0241809.
XX XX 20-OCT-2000; 2000US-0241826.
XX XX 01-NOV-2000; 2000US-0244617.
XX XX 08-NOV-2000; 2000US-0246474.
XX XX 08-NOV-2000; 2000US-0246475.
XX XX 08-NOV-2000; 2000US-0246476.
XX XX 08-NOV-2000; 2000US-0246477.
XX XX 08-NOV-2000; 2000US-0246478.
XX XX 08-NOV-2000; 2000US-0246523.
XX XX 08-NOV-2000; 2000US-0246524.
XX XX 08-NOV-2000; 2000US-0246525.
XX XX 08-NOV-2000; 2000US-0246526.
XX XX 08-NOV-2000; 2000US-0246527.
XX XX 08-NOV-2000; 2000US-0246528.
XX XX 08-NOV-2000; 2000US-0246532.
XX XX 08-NOV-2000; 2000US-0246609.
XX XX 08-NOV-2000; 2000US-0246610.
XX XX 08-NOV-2000; 2000US-0246611.
XX XX 08-NOV-2000; 2000US-0246613.
XX XX 17-NOV-2000; 2000US-0249207.
XX XX 17-NOV-2000; 2000US-0249208.
XX XX 17-NOV-2000; 2000US-0249209.
XX XX 17-NOV-2000; 2000US-0249210.
XX XX 17-NOV-2000; 2000US-0249211.
XX XX 17-NOV-2000; 2000US-0249212.
XX XX 17-NOV-2000; 2000US-0249213.
XX XX 17-NOV-2000; 2000US-0249214.
XX XX 17-NOV-2000; 2000US-0249215.
XX XX 17-NOV-2000; 2000US-0249216.
XX XX 17-NOV-2000; 2000US-0249217.
XX XX 17-NOV-2000; 2000US-0249218.
XX XX 17-NOV-2000; 2000US-0249244.
XX XX 17-NOV-2000; 2000US-0249245.
XX XX 17-NOV-2000; 2000US-0249264.
XX XX 17-NOV-2000; 2000US-0249265.
XX XX 17-NOV-2000; 2000US-0249297.
XX XX 17-NOV-2000; 2000US-0249299.
XX XX 17-NOV-2000; 2000US-0249300.
XX XX 01-DEC-2000; 2000US-0250160.
XX XX 01-DEC-2000; 2000US-0250391.
XX XX 05-DEC-2000; 2000US-0251030.
XX XX 05-DEC-2000; 2000US-0251988.

PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 08-NOV-2000; 2000US-0246617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-02559678.
 (HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Barash SC, Ruben SM;

PI WPI; 2001-483426/52.

DR Nucleic acids encoding human immune/hematopoietic antigen polypeptides,

XX useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 PT

XX

PS Disclosure; SEQ ID NO 29132; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 5572 BP; 1315 A; 1312 C; 1444 G; 1501 T; 0 other;
 Query Match 5.0%; Score 50; DB 22; Length 5572;
 Best Local Similarity 100.0%; Pred. No. 9.7e-14;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 338 TCTACTAAATAACAAAAATTAGCCAGGTGTGGTGGCACACGCTGTAGT 387
 Db 4009 TCTACTAAATAACAAAAATTAGCCAGGTGTGGTGGCACACGCTGTAGT 4058
 RESULT 40
 AAK74321
 ID AAK74321 standard; DNA; 5573 BP.
 AC AAK74321;
 XX
 XX 07-NOV-2001 (first entry)
 DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:29133.
 XX
 KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 XX cytostatic; gene therapy; vaccine; metastasis; ds.
 OS Homo sapiens.
 XX
 PN WO200157182-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01354.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 28-JUN-2000; 2000US-0209467.
 PR 30-JUN-2000; 2000US-0214886.
 PR 07-JUL-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 11-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 14-JUL-2000; 2000US-0217496.
 PR 26-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 14-AUG-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-02252114.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226686.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0234998.
 PR 27-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239335.
 PR 13-OCT-2000; 2000US-0239337.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0245474.
 PR 08-NOV-2000; 2000US-0245475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SK;
 XX
 XX WPI; 2001-483426/52.
 DR
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 PS Disclosure; SEQ ID NO 29133; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK62169
 CC represent sequences used in the exemplification of the present invention.
 XX

SQ Sequence 5573 BP; 1314 A; 1312 C; 1446 G; 1501 T; 0 other;
Query Match 5.0%; Score 50; DB 22; Length 5573;
Best Local Similarity 100.0%; Pred. No. 9.7e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 338 TCTACTAAATAACAAAATAGCCAGGTGCTGGTGCACAGCCTGTAGT 387
|||||
Db 4010 TCTACTAAATAACAAAATAGCCAGGTGCTGGTGCACAGCCTGTAGT 4059
RESULT 41
AAS46383
ID AAS46383 standard; DNA; 5649 BP.
XX AAS46383;
AC AAS46383;
XX
DT 18-DEC-2001 (first entry)
XX
DE Tumour suppressor gene derived chemically modified sequence #105.
XX
KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW cytosine methylation; ds.
XX
OS Homo sapiens.
XX
PN WO200168912-A2.
XX
PD 20-SEP-2001.
XX
PF 15-MAR-2001; 2001WO-EP02955.
XX
PR 15-MAR-2000; 2000DE-1013847.
PR 06-APR-2000; 2000DE-1019058.
PR 07-APR-2000; 2000DE-1019173.
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-602752/68.
XX
PT Fragments of chemically modified genes associated with tumour suppressor
PT genes and oncogenes, useful in designing primers and probes for
PT analysing diseases associated with cytosine methylation state e.g.
PT cancer -
XX
PS Claim 1; SEQ ID No 105; 27pp; English.
XX
CC The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and
CC oncogenes having a sequence taken from 536 (actually 533 since
CC (SS) and sequences complementary to (SS). The nucleic acid may be a
CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
CC form part of a set of probes for detecting the cytosine methylation state
CC array for analysing diseases associated with CpG dinucleotides e.g.
CC cancers and tumours. The probes can also be used in a method for
CC ascertaining genetic and/or epigenetic parameters for the diagnosis
CC and/or therapy of existing diseases or the predisposition to specific
CC diseases, by analysing cytosine methylations. The parameters may be
CC compared to another set of genetic and/or epigenetic parameters, the
CC differences serving as basis for diagnosis and/or prognosis events which
CC are disadvantageous to patients. The present sequence is one of the
CC 533 genomic sequences derived from tumour suppressor genes and
CC oncogenes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic

CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 5649 BP; 1566 A; 184 C; 1474 G; 2425 T; 0 other;
Query Match 5.0%; Score 50; DB 22; Length 5649;
Best Local Similarity 100.0%; Pred. No. 9.7e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 509 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 558
|||||
Db 4508 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 4557
RESULT 42
AAS46384/C
ID AAS46384 standard; DNA; 5649 BP.
XX
XX AAS46384;
AC AAS46384;
XX
DT 18-DEC-2001 (first entry)
XX
DE Tumour suppressor gene derived chemically modified sequence #106.
XX
KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW cytosine methylation; ds.
XX
OS Homo sapiens.
XX
PN WO200168912-A2.
XX
PD 20-SEP-2001.
XX
PF 15-MAR-2001; 2001WO-EP02955.
XX
PR 15-MAR-2000; 2000DE-1013847.
PR 06-APR-2000; 2000DE-1019058.
PR 07-APR-2000; 2000DE-1019173.
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-602752/68.
XX
PT Fragments of chemically modified genes associated with tumour suppressor
PT genes and oncogenes, useful in designing primers and probes for
PT analysing diseases associated with cytosine methylation state e.g.
PT cancer -
XX
PS Claim 1; SEQ ID No 106; 27pp; English.
XX
CC The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and
CC oncogenes having a sequence taken from 536 (actually 533 since
CC (SS) and sequences complementary to (SS). The nucleic acid may be a
CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
CC form part of a set of probes for detecting the cytosine methylation state
CC array for analysing diseases associated with CpG dinucleotides e.g.
CC cancers and tumours. The probes can also be used in a method for
CC ascertaining genetic and/or epigenetic parameters for the diagnosis
CC and/or therapy of existing diseases or the predisposition to specific
CC diseases, by analysing cytosine methylations. The parameters may be
CC compared to another set of genetic and/or epigenetic parameters, the
CC differences serving as basis for diagnosis and/or prognosis events which
CC are disadvantageous to patients. The present sequence is one of the
CC 533 genomic sequences derived from tumour suppressor genes and

CC oncogenes. Sequences with even numbered Seq ID numbers are the
CC complementary sequence of the corresponding odd numbered sequence (e.g.
CC ID 2 and ID1, ID 536 and ID 535, except for those whose partner sequence
CC is missing).
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX

XX Seq Sequence 5649 BP; 1212 A; 184 C; 1397 G; 2856 T; 0 other;
Query Match 5.0%; Score 50; DB 22; Length 5649;
Best Local Similarity 100.0%; Pred. No. 9.7e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 509 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 558
|||||
Db 1142 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 1093
|||||

RESULT 43
ABK40007
ID ABK40007 standard; DNA; 5649 BP.
XX
AC ABK40007;
XX
DT 21-MAY-2002 (first entry)
XX
DE Human chemically pretreated gene sequence #45 strand 1.
XX

XX Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;
KW Cytostatic; ALDH6; CYP11A; CYP11B1; CYP3A3; DPYD; EPHX2; OCLN; TXNRD1;
KW UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.

XX WO200202806-A2.
XX
PN 10-JAN-2002.
XX
PD 29-JUN-2001; 2001WO-EF07470.
XX
PF 30-JUN-2000; 2000DE-1032529.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-154757/20.
XX
DR New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,
XX useful for detecting cytosine methylation state of genes associated
XX with pharmacogenomics and for therapy of diseases e.g. cancer
XX
PS Claim 1; SEQ ID No 89; 24pp; English.

XX The invention relates to a nucleic acid comprising a sequence at
XX least 18 bases in length of a segment of the chemically pretreated DNA
XX of genes associated with pharmacogenomics according to one of the
XX sequences of the genes ALDH6 (NM_000693), CYP11A (NM_000781), CYP11B1
XX (NM_000497), CYP3A3 (NM_000776 and NM_017460), DPYD (NM_000110), EPHX2
XX (NM_001979), OCLN (NM_002538), TXNRD1 (NM_003330), UGT8 (NM_003360),
XX MRP (NM_004996, NM_019900, NM_019901, NM_019902, NM_019862, NM_019898,
XX NM_019899) and their complementary sequences, or a sequence (S1) chosen
XX from 87 sequences and their complements. The chemical pretreatment
XX is bisulphite treatment to convert cytosines (but not methyl-cytosines)
XX into uracils. Also included are an oligomer (II) in particular an
XX oligonucleotide or a peptide nucleic acid (PNA)-oligomer, comprising in
XX each case at least one base sequence having a length of 9 nucleotides
XX which hybridises to or is identical to a chemically pretreated DNA of
XX genes associated with pharmacogenomics and their complements, arranged in
XX an array for analysing diseases associated with the methylation state

CC (CpG) and/or detecting SNPs (single nucleotide polymorphisms)
CC of the 87 sequences. The oligomers may also be used as PCR primers.
CC The set of 87 nucleic acids and their complements is useful for diagnosis
CC and therapy of solid tumours and cancer. The present sequence
CC represents one the 87 DNA sequences or its complement.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX

XX Seq Sequence 5649 BP; 1566 A; 184 C; 1474 G; 2425 T; 0 other;
Query Match 5.0%; Score 50; DB 24; Length 5649;
Best Local Similarity 100.0%; Pred. No. 9.7e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 509 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 558
|||||
Db 4508 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 4557
|||||

RESULT 44
ABK40008/c
ID ABK40008 standard; DNA; 5649 BP.
XX
AC ABK40008;
XX
DT 21-MAY-2002 (first entry)
XX
DE Human chemically pretreated gene sequence #45 strand 2.

XX Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;
KW Cytostatic; ALDH6; CYP11A; CYP11B1; CYP3A3; DPYD; EPHX2; OCLN; TXNRD1;
KW UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.

XX WO200202806-A2.
XX
PN 10-JAN-2002.
XX
PD 29-JUN-2001; 2001WO-EF07470.
XX
PF 30-JUN-2000; 2000DE-1032529.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-154757/20.
XX
DR New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,
XX useful for detecting cytosine methylation state of genes associated
XX with pharmacogenomics and for therapy of diseases e.g. cancer
XX
PS Claim 1; SEQ ID No 90; 24pp; English.

XX The invention relates to a nucleic acid comprising a sequence at
XX least 18 bases in length of a segment of the chemically pretreated DNA
XX of genes associated with pharmacogenomics according to one of the
XX sequences of the genes ALDH6 (NM_000693), CYP11A (NM_000781), CYP11B1
XX (NM_000497), CYP3A3 (NM_000776 and NM_017460), DPYD (NM_000110), EPHX2
XX (NM_001979), OCLN (NM_002538), TXNRD1 (NM_003330), UGT8 (NM_003360),
XX MRP (NM_004996, NM_019900, NM_019901, NM_019902, NM_019862, NM_019898,
XX NM_019899) and their complementary sequences, or a sequence (S1) chosen
XX from 87 sequences and their complements. The chemical pretreatment
XX is bisulphite treatment to convert cytosines (but not methyl-cytosines)
XX into uracils. Also included are an oligomer (II) in particular an
XX oligonucleotide or a peptide nucleic acid (PNA)-oligomer, comprising in
XX each case at least one base sequence having a length of 9 nucleotides
XX which hybridises to or is identical to a chemically pretreated DNA of
XX genes associated with pharmacogenomics and their complements, arranged in

CC an array for analysing diseases associated with the methylation state
CC (CpG) and/or detecting SNPs (single nucleotide polymorphisms)
CC of the 87 sequences. The oligomers may also be used as PCR primers.
CC The set of 87 nucleic acids and their complements is useful for diagnosis
CC and therapy of solid tumours and cancer. The present sequence
CC represents one the 87 DNA sequences or its complement.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 5649 BP; 1212 A; 184 C; 1397 G; 2856 T; 0 other;

Query Match 5.0%; Score 50; DB 24; Length 5649;
Best Local Similarity 100.0%; Pred. No. 9.7e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 509 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 558
|||||
Db 1142 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 1093

RESULT 45

ABL32848

ID ABL32848 standard; DNA; 5649 BP.

XX AC ABL32848;

XX

DT 26-MAR-2002 (first entry)

XX

DE Human immune system associated gene SEQ ID NO: 821.

XX

KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriasis;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.

XX OS Homo sapiens.

XX

PN WO200200928-A2.

XX

PD 03-JAN-2002.

XX

PF 02-JUL-2001; 2001WO-EP07537.

XX

PR 30-JUN-2000; 2000DE-1032529.

XX

PR 01-SEP-2000; 2000DE-1043826.

XX

PA (EPig-) EPIGENOMICS AG.

XX

PI Olek A, Piepenbrock C, Berlin K;

XX

DR WPI; 2002-130909/17.

XX

PT Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation

XX

PS Claim 1; SEQ ID NO 821; 32pp + Sequence Listing; German.

XX

CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases, such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.

XX

SQ Sequence 5649 BP; 1566 A; 184 C; 1474 G; 2425 T; 0 other;
Query Match 5.0%; Score 50; DB 24; Length 5649;
Best Local Similarity 100.0%; Pred. No. 9.7e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 509 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 558
|||||
Db 4508 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 4557

Search completed: November 24, 2002, 10:51:00
Job time : 1558 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2002, 09:22:22 ; Search time 3223 Seconds
(without alignments)
9029.726 Million cell updates/sec

Title: US-09-784-423-32
Perfect score: 1000
Sequence: 1 GTGTGACCTTATCTCTCT.....CAGATGATGACGGCGTGC 1000

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.in.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pl.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.mu.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg_hum.*

31: em.htg_inv.*

32: em.htg_other.*

33: em.htg_mus.*

34: em.htg_pin.*

35: em.htg_rod.*

36: em.htg_mam.*

37: em.htg_vrt.*

38: em.sy.*

39: em.htgo_hum.*

40: em.htgo_mus.*

41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Match	Length	DB	ID	Description
c	1	1000	100.0	1000	6	ARI53982	ARI53982 Sequence
c	2	879	87.9	69964	9	HS496C20	Z83847 Human DNA s
c	3	492	49.2	138459	2	AC090795	AC090795 Homo sapi
c	4	304	30.4	62102	2	AC104363	AC104363 Homo sapi
c	5	77	7.7	102151	9	HS979X1	AL035659 Human DNA
c	6	76	7.6	166869	9	AC016395	AC016395 Homo sapi
c	7	76	7.6	170998	9	AL390071	AL390071 Human DNA
c	8	76	7.6	290625	2	AL442068	AL442068 Homo sapi
c	9	75	7.5	59427	9	AC079882	AC079882 Homo sapi
c	10	75	7.5	124821	9	AC004983	AC004983 Homo sapi
c	11	74	7.4	167417	9	AC011505	AC011505 Homo sapi
c	12	72	7.2	96308	9	AF190465	AF190465 Homo sapi
c	13	72	7.2	173354	9	HS149A16	AL021937 Human DNA
c	14	72	7.2	273403	9	AC011498	AC011498 Homo sapi
c	15	71	7.1	162201	2	AC024152	AC024152 Homo sapi
c	16	71	7.1	186241	9	AC103560	AC103560 Homo sapi
c	17	70	7.0	163396	9	AL161445	AL161445 Human DNA
c	18	70	7.0	197143	2	AC069173	AC069173 Homo sapi
c	19	67	6.7	170513	2	AC023603	AC023603 Homo sapi
c	20	67	6.7	178361	9	AC006042	AC006042 Homo sapi
c	21	66	6.6	104656	9	AL731576	AL731576 Human DNA
c	22	66	6.6	121028	9	AY052369	AY052369 Homo sapi
c	23	66	6.6	166975	9	AC110023	AC110023 Homo sapi
c	24	66	6.6	182627	2	AC024262	AC024262 Homo sapi
c	25	66	6.6	182776	9	AC091078	AC091078 Homo sapi
c	26	66	6.6	186431	2	AC022281	AC022281 Homo sapi
c	27	66	6.6	189116	2	CNS01DWM	AL137779 Human chr
c	28	64	6.4	127140	9	AL136137	AL136137 Human DNA
c	29	64	6.4	146954	9	AL590640	AL590640 Human DNA
c	30	64	6.4	173510	9	AL359272	AL359272 Homo sapi
c	31	64	6.4	177242	2	AC021923	AC021923 Homo sapi
c	32	64	6.4	187349	9	AP003352	AP003352 Homo sapi
c	33	64	6.4	197870	2	AC091075	AC091075 Homo sapi
c	34	64	6.4	199208	2	AC019236	AC019236 Homo sapi
c	35	63	6.3	62443	9	AL136973	AL136973 Human DNA
c	36	63	6.3	85713	9	HSY237C10_3	Continuation (4 of
c	37	63	6.3	160010	2	AL844148	AL844148 Homo sapi
c	38	63	6.3	168071	2	AL391479	AL391479 Homo sapi
c	39	63	6.3	182789	9	AC021761	AC021761 Homo sapi
c	40	63	6.3	212730	2	AL354975	AL354975 Homo sapi
c	41	63	6.3	232162	2	AL591130	AL591130 Homo sapi
c	42	62	6.2	68122	9	AP000998	AP000998 Homo sapi
c	43	62	6.2	71530	9	AC025810	AC025810 Homo sapi
c	44	62	6.2	124211	9	AC016593	AC016593 Homo sapi
c	45	62	6.2	143146	9	AC079855	AC079855 Homo sapi

ALIGNMENTS

RESULT 1	ARI53982	1000 bp	DNA	linear	PAT 08-AUG-2001
LOCUS	Sequence 32 from patent US 6238863.				
DEFINITION	ARI53982				
ACCESSION	ARI53982.1	GI:15122035			
VERSION					
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified				
REFERENCE	1 (bases 1 to 1000)				
AUTHORS	Schumm,J.W. and Bachter,J.W.				
TITLE	Materials and methods for indentifying and analyzing intermediate tandem repeat DNA markers				
JOURNAL	Patent: US 6238863-A 32 29-MAY-2001;				

```
FEATURES
Source      Location/Qualifiers
BASE COUNT  302 a 240 c 231 g 226 t 1 others
ORIGIN
Query Match      100.0%; Score 1000; DB 6; Length 1000;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTGTGACCTTATCTCTCTGAACCTCAGTTTCCTTCATCCGTAATGAAAGTGGCTAG 60
Db 1 GGTGTGACCTTATCTCTCTGAACCTCAGTTTCCTTCATCCGTAATGAAAGTGGCTAG 60
QY 61 ATTGCTTAAATAAATAATTAATGGAATAGGCTAGCGGGTGGCTCAGCGCTGTAATCCCA 120
Db 61 ATTGCTTAAATAAATAATTAATGGAATAGGCTAGCGGGTGGCTCAGCGCTGTAATCCCA 120
QY 121 GCACCTTTAGAGGTGGAAGAGGTGGATCAGTTCAGTTCAGGAGTTCCTGAGACGAGCTG 180
Db 121 GCACCTTTAGAGGTGGAAGAGGTGGATCAGTTCAGTTCAGGAGTTCCTGAGACGAGCTG 180
QY 181 GCCAACAGGTGAAACCCCATCTCTACTTAAATAAATAAATAATAGCTGCGGTGGCT 240
Db 181 GCCAACAGGTGAAACCCCATCTCTACTTAAATAAATAAATAATAGCTGCGGTGGCT 240
QY 241 CACACCTGTAATCCAGCAGCTTTGGAGGCTGAGACGGGTGGATCAGTTCAGTTCAGGAG 300
Db 241 CACACCTGTAATCCAGCAGCTTTGGAGGCTGAGACGGGTGGATCAGTTCAGTTCAGGAG 300
QY 301 TTCAGGCGAGGCTGGGCAACATGGTGAACACCGTCTCTACTTAAATAAATAAATAATTA 360
Db 301 TTCAGGCGAGGCTGGGCAACATGGTGAACACCGTCTCTACTTAAATAAATAAATAATTA 360
QY 361 CCAGGTGTGGTGACACAGGCTGTAGTCCAGCTTACTTGGAGGCTGAGGCGGAAGAATC 420
Db 361 CCAGGTGTGGTGACACAGGCTGTAGTCCAGCTTACTTGGAGGCTGAGGCGGAAGAATC 420
QY 421 GCTTGAACCCAGTAGGAGGCTGGAGGCTGAGCGGAGATGAAGTCACTGCTCCAGCC 480
Db 421 GCTTGAACCCAGTAGGAGGCTGGAGGCTGAGCGGAGATGAAGTCACTGCTCCAGCC 480
QY 481 TGGGTGACAGAGCAAGACTCCCTCTCAGAAATAAATAAATAAATAAATAAATAAATAA 540
Db 481 TGGGTGACAGAGCAAGACTCCCTCTCAGAAATAAATAAATAAATAAATAAATAAATAA 540
QY 541 AATAAATAAATAAATAAATAAAGGCTGGCATTTGGCTAGCACTTATATGCCCAATA 600
Db 541 AATAAATAAATAAATAAATAAAGGCTGGCATTTGGCTAGCACTTATATGCCCAATA 600
QY 601 AGTAATAGCTATCAATATCCACACCCCTACCACTGTGCTGAATTTAGTTCTTTTGTG 660
Db 601 AGTAATAGCTATCAATATCCACACCCCTACCACTGTGCTGAATTTAGTTCTTTTGTG 660
QY 661 ACCGCCATTAAGCTTAAGCGAGAATCTCACCGTACTCTCTGTAAATTTCTGTTCT 720
Db 661 ACCGCCATTAAGCTTAAGCGAGAATCTCACCGTACTCTCTGTAAATTTCTGTTCT 720
QY 721 GGCACATAGTTGGTCTCAGTGAACATGTTGAGTGAATGAGCAAAATGCAAGAAATCTCC 780
Db 721 GGCACATAGTTGGTCTCAGTGAACATGTTGAGTGAATGAGCAAAATGCAAGAAATCTCC 780
QY 781 AGGCCATCTGGAGGCCCTCCAGGCGGCTGAGTTCGGGAACTCATAGTCTGCTCTCAAT 840
Db 781 AGGCCATCTGGAGGCCCTCCAGGCGGCTGAGTTCGGGAACTCATAGTCTGCTCTCAAT 840
QY 841 GGCCCACTGAAGGTAGAGAGTTCGGGTCCCACTCCGACCCCACTCTCTCTGACTCAC 900
Db 841 GGCCCACTGAAGGTAGAGAGTTCGGGTCCCACTCCGACCCCACTCTCTCTGACTCAC 900
QY 901 TGCTGAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 960
Db 901 TGCTGAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 960
```

```
QY 961 CAGGACTCCAGGAGCCACACAGATGATGACCGGCGTGC 1000
Db 961 CAGGACTCCAGGAGCCACACAGATGATGACCGGCGTGC 1000

RESULT 2
HS496C20/c
LOCUS
DEFINITION
  Human DNA sequence from clone RP3-496C20 on chromosome 22 Contains
  STSs, GSSs and two putative CpG islands, complete sequence.
VERSION
  Z83847.1 GI:3550060
KEYWORDS
  HTG; CpG island.
SOURCE
  Homo sapiens.
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  Clark.G.
  Direct Submission
  Submitted (08-DEC-1999) Sanger Centre, Hinxton, Cambridgeshire,
  CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
  Requests: clonerequest@sanger.ac.uk
  On Sep 8, 1998 this sequence version replaced gi:3413284.
  This sequence has been finished according to sequence map criteria
  as follows. An attempt is made to resolve all sequencing problems,
  such as compressions and repeats, but not necessarily within known
  annotated human repeat sequence elements (e.g. Alu). Where the
  sequence is ambiguous, there is an annotation using the 'unsure'
  feature key.
  This sequence was generated from part of bacterial clone contigs of
  human chromosome 22, constructed by the Sanger Centre Chromosome 22
  Mapping Group. Further information can be found at
  http://www.sanger.ac.uk/HGP/Chr22
  During sequence assembly data is compared from overlapping clones.
  Where differences are found these are annotated as variations
  together with a note of the overlapping clone name. Note that the
  variation annotation may not be found in the sequence submission
  corresponding to the overlapping clone, as we submit sequences with
  only a small overlap as described above.
  The following abbreviations are used to associate primary accession
  numbers given in the feature table with their source databases:
  Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information
  on the WORMPEP database can be found at
  http://www.sanger.ac.uk/Projects/C/elegans/wormpep
  from the library RPI-3 constructed at the Roswell Park Cancer
  Institute by the group of Pieter de Jong. For further details see
  http://bacpac.med.buffalo.edu/
  VECTOR: pCYPAC2
  This sequence is the entire insert of clone RP3-496C20. The true
  left end of clone RP3-340K22 is at 21972 in this sequence. The true
  right end of clone RPI-140N12 is at 41206 in this sequence. The
  start of this sequence overlaps with sequence Z82206. The end of
  this sequence overlaps with sequence AL022238.
  Location/Qualifiers
    1..69964
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /chromosome="22"
      /clone="RP3-496C20"
      /clone_lib="RPI-3"
      /note="match: STS: Em:HS496C20S; match: STS: Em:AL021673"
      169..311
        /note="MIR repeat: matches 71..211 of consensus"
        551..600
          /note="10 copies 5 mer tgttt 76 conserved"
          2394..2462
            /note="L2 repeat: matches 2626..2696 of consensus"
            3060..3221
              /note="MIR repeat: matches 21..192 of consensus"
              3226..3535

FEATURES
source
  1..69964
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="22"
    /clone="RP3-496C20"
    /clone_lib="RPI-3"
    /note="match: STS: Em:HS496C20S; match: STS: Em:AL021673"
    169..311
      /note="MIR repeat: matches 71..211 of consensus"
      551..600
        /note="10 copies 5 mer tgttt 76 conserved"
        2394..2462
          /note="L2 repeat: matches 2626..2696 of consensus"
          3060..3221
            /note="MIR repeat: matches 21..192 of consensus"
            3226..3535

misc_feature
  repeat_region
  repeat_region
  repeat_region
  repeat_region
  repeat_region
```

```
repeat_region /note="AluJb repeat: matches 5. .312 of consensus"
repeat_region /note="MIR repeat: matches 35. .236 of consensus"
misc_feature 5924. .5970
/note="L2 repeat: matches 2701. .2748 of consensus"
6015. .6330
/note="match: GSS: Em:AQ356083"
6149. .6461
/note="Alusx repeat: matches 1. .312 of consensus"
6462. .6596
/note="Alusq/x repeat: matches 1. .135 of consensus"
6632. .6695
/note="MIR repeat: matches 70. .146 of consensus"
7817. .8113
/note="Alusx repeat: matches 1. .297 of consensus"
complement(8207. .8891)
/note="match: GSS: Em:AQ344529"
8840. .8926
/note="MIR repeat: matches 80. .191 of consensus"
9079. .9606
/note="264 copies 2 mer cc 58 conserved"
9187. .9189
/note="forced join; gap estimated as 1kb by restriction
digest data"
complement(9187. .9189)
/note="Forced join. Gap estimated as 1kb by restriction
digest data."
9481. .9605
/note="25 copies 5 mer ctct 62 conserved"
10240. .10395
/note="MIR repeat: matches 76. .255 of consensus"
10502. .10587
/note="MER5A repeat: matches 57. .150 of consensus"
10511. .10643
/note="MER5A repeat: matches 57. .185 of consensus"
11303. .11539
/note="L2 repeat: matches 1865. .2127 of consensus"
13046. .13314
/note="Alusx repeat: matches 28. .305 of consensus"
13527. .13983
/note="L1M4 repeat: matches 5393. .5854 of consensus"
13984. .14293
/note="Alusq repeat: matches 1. .310 of consensus"
14294. .14347
/note="L1M48 repeat: matches 5340. .5393 of consensus"
14348. .14649
/note="Alusx repeat: matches 2. .304 of consensus"
14651. .14966
/note="Aluy repeat: matches 1. .306 of consensus"
14967. .15007
/note="L1M48 repeat: matches 5352. .5341 of consensus"
15024. .16371
/note="L1M4 repeat: matches 3784. .5097 of consensus"
16413. .16611
/note="L1 repeat: matches 3186. .3382 of consensus"
16662. .16875
/note="MIR repeat: matches 29. .256 of consensus"
16891. .17310
/note="MLT2FB repeat: matches 2. .414 of consensus"
17137. .17686
/note="match: GSS: Em:AQ551353"
17359. .17438
/note="MER31A repeat: matches 114. .193 of consensus"
17825. .18044
/note="LTR23 repeat: matches 210. .437 of consensus"
18046. .18357
/note="Alusq repeat: matches 1. .309 of consensus"
18511. .18822
/note="Alusq repeat: matches 1. .313 of consensus"
18830. .18897
/note="BURL repeat: matches 6300. .6367 of consensus"
19044. .19080
/note="MIR repeat: matches 207. .245 of consensus"
```

```
repeat_region 19081. .19391
/note="Alusq repeat: matches 1. .307 of consensus"
19392. .19492
/note="MIR repeat: matches 117. .207 of consensus"
19581. .19801
/note="MIR repeat: matches 29. .261 of consensus"
20302. .20464
/note="Alusq/x repeat: matches 134. .302 of consensus"
20465. .20747
/note="Alusq repeat: matches 1. .296 of consensus"
22217. .22508
/note="Alusx repeat: matches 14. .305 of consensus"
22764. .22848
/note="MIR repeat: matches 164. .249 of consensus"
23209. .23725
/note="CpG island"
23946. .24063
/evidence=not_experimental
/note="MIR repeat: matches 21. .140 of consensus"
25053. .25142
/note="MIR repeat: matches 21. .119 of consensus"
25207. .25233
/note="9 copies 3 mer gct 96 conserved"
25401. .26038
/note="match: GSS: Em:B54652"
25877. .26155
/note="AluJb repeat: matches 1. .296 of consensus"
26997. .27304
/note="AluJo repeat: matches 1. .311 of consensus"
27326. .27595
/note="Alusx repeat: matches 37. .311 of consensus"
28098. .28406
/note="Alusq repeat: matches 1. .311 of consensus"
28847. .29128
/note="L2 repeat: matches 2362. .2745 of consensus"
complement(28031. .29545)
/note="match: GSS: Em:AQ528941"
30940. .30989
/note="25 copies 2 mer gt 100 conserved"
30990. .31054
/note="MER41A repeat: matches 400. .464 of consensus"
31120. .31339
/note="match: STS: Em:G03731; match: STS: Em:G03731"
34416. .34721
/note="Aluy repeat: matches 1. .310 of consensus"
34776. .34821
/note="L2 repeat: matches 2688. .2733 of consensus"
34944. .35243
/note="Alusx repeat: matches 1. .300 of consensus"
35351. .35551
/note="MER20 repeat: matches 1. .204 of consensus"
36460. .36544
/note="L2 repeat: matches 2665. .2749 of consensus"
37480. .37620
/note="MIR repeat: matches 4. .150 of consensus"
37861. .38188
/note="AluJo repeat: matches 1. .312 of consensus"
complement(38024. .38419)
/note="match: STS: Em:G28177; match: STS: Em:R54672"
complement(38086. .38419)
/note="match: STS: Em:G25989"
38802. .38948
```

Query Match 87.9%; Score 879; DB 9; Length 69984;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 999; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 GGTCGACCTTATCTCTGACCTCAGTTTCTCATCCGTAATAATGAAAGCTGTAG 60
|||||
Db 6893 GGTCGACCTTATCTCTGACCTCAGTTTCTCATCCGTAATAATGAAAGCTGTAG 6624
|||||
Qy 61 ATTGTTGTAATAAATAATGAATAGGTAGCGGGTGGCTCACCCCTGTAATCCCA 120
|||||

Db	6623	ATTGTTGTAATAAATAAATGAATAGGCTAGCGCGTGGCTCAGCGCTGTAATCCCA	6564
QY	121	GCACCTTTAGAGGTGCAAGAGGGTGGATCACTTGAGGTGAGGAGTTTGGAGACGAGCTG	180
Db	6563	GCACCTTTAGAGGTGCAAGAGGGTGGATCACTTGAGGTGAGGAGTTTGGAGACGAGCTG	6504
QY	181	GCACACGCTGAAACCCCTCTCTACTAAATAAATAAATAAATAAATAAATAAATAAATAA	240
Db	6503	GCACACGCTGAAACCCCTCTCTACTAAATAAATAAATAAATAAATAAATAAATAAATAA	6445
QY	241	CACACCTGTAATCCCGACCTTTGGGAGCGTGGAGCGGTGGATCACTTGAAGTCAGGAG	300
Db	6444	CACACCTGTAATCCCGACCTTTGGGAGCGTGGAGCGGTGGATCACTTGAAGTCAGGAG	6385
QY	301	TTCAAGCGCAGCTGGCGACATGCTGAACACAGTCTTACTTAAATAAATAAATAAATAA	360
Db	6384	TTCAAGCGCAGCTGGCGACATGCTGAACACAGTCTTACTTAAATAAATAAATAAATAA	6325
QY	361	CCAGGTGTTGGTGGCACACGCGCTGTAGTCCCGACTACTTGGAGGCTGAGGCGGAAGATC	420
Db	6324	CCAGGTGTTGGTGGCACACGCGCTGTAGTCCCGACTACTTGGAGGCTGAGGCGGAAGATC	6265
QY	421	GCTTGAACCCAGTAGGAGAGTGGCGAGTGGAGCGGAGATGAAGTCACTGCCTCCAGCC	480
Db	6264	GCTTGAACCCAGTAGGAGAGTGGCGAGTGGAGCGGAGATGAAGTCACTGCCTCCAGCC	6205
QY	481	TGGTGACAGACGAGACTCCCTCTCAGAAAATAAATAAATAAATAAATAAATAAATAA	540
Db	6204	TGGTGACAGACGAGACTCCCTCTCAGAAAATAAATAAATAAATAAATAAATAAATAA	6145
QY	541	AATAAATAAATAAATAAATAAAGGCTGGCATTTGGCTAGCATTATATGCCAATA	600
Db	6144	AATAAATAAATAAATAAATAAAGGCTGGCATTTGGCTAGCATTATATGCCAATA	6085
QY	601	AGTAATAGCTATCAATATCCCAACCGCTACCACTGTGCTGAAATTTAGTTCTTTGTC	660
Db	6084	AGTAATAGCTATCAATATCCCAACCGCTACCACTGTGCTGAAATTTAGTTCTTTGTC	6025
QY	661	ACCCGCCATTAGCTTAAGCGAGATTTCTACCGTACTCTCTGTAATTTCTGTTCT	720
Db	6024	ACCCGCCATTAGCTTAAGCGAGATTTCTACCGTACTCTCTGTAATTTCTGTTCT	5965
QY	721	GGCACAATGGTGTCTCAGTGAACATGTTGAGTGAATGAGCAAAATGCAAGAACTCC	780
Db	5964	GGCACAATGGTGTCTCAGTGAACATGTTGAGTGAATGAGCAAAATGCAAGAACTCC	5905
QY	781	AGCCATCTGGAGCCCTCCAGCGGGTGGTGGGAACTCATAGTCTGTCTCAAT	840
Db	5904	AGCCATCTGGAGCCCTCCAGCGGGTGGTGGGAACTCATAGTCTGTCTCAAT	5845
QY	841	GGCCCACTGAAAGTGAAGTCTTGGGTCCACCTCCGACCGCCCTCTCTGACTCAC	900
Db	5844	GGCCCACTGAAAGTGAAGTCTTGGGTCCACCTCCGACCGCCCTCTCTGACTCAC	5785
QY	901	TGCTGAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	960
Db	5784	TGCTGAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	5725
QY	961	CAGGACTCAGGAGCCCGACGAGATGATGACCGGGTGC	1000
Db	5724	CAGGACTCAGGAGCCCGACGAGATGATGACCGGGTGC	5685
RESULT	3		
LOCUS	AC090795	138459 bp	DNA linear
DEFINITION	Homo sapiens chromosome 8 clone RP11-254K5 map 8, WORKING DRAFT		
ACCESSION	AC090795		
VERSION	AC090795.2	GI:14336499	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULITOP.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 138459) Birren, B., Linton, L., Nusbaum, C. and Lander, E. Homo sapiens chromosome 8, clone RP11-254K5 Unpublished 2 (bases 1 to 138459) Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A., Canarita, J., Campopiano, A., Choepel, Y., Colangelo, M., Collins, S., Conlaymore, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Govette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatas, A., LaRocque, K., Lamazares, R., Landers, T., Lehotzky, J., Levine, K., Liu, G., Maclean, C., Macdonald, P., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., Mihova, T., Mlekena, K., McPheeters, R., Meldrum, J., Meneus, L., Mihova, T., Mlekena, K., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Sougnez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M. Direct Submission Submitted (10-MAR-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA On Jun 11, 2001 this sequence version replaced gi:13270664. All repeats were identified using RepeatMasker: Smit, A. F. A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html ----- Genome Center Center: Whitehead Institute/ MIT Center for Genome Research Center code: WIBR Web site: http://www-seq.wi.mit.edu Contact: sequence_submissions@genome.wi.mit.edu ----- Project Information Center project name: L12790 Center clone name: 254_K_5 ----- Summary Statistics Sequencing vector: Plasmid; n/a; 100% of reads Chemistry: Dye-terminator Big Dye; 100% of reads Assembly program: Phrap; version 0.960731 Consensus quality: 133567 bases at least Q40 Consensus quality: 135811 bases at least Q30 Consensus quality: 136678 bases at least Q20 Insert size: 139000; agarose-fp Insert size: 137359; sum-of-contigs Quality coverage: 6.5 in Q20 bases; agarose-fp Quality coverage: 6.6 in Q20 bases; sum-of-contigs ----- * NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved. * 1 1411: contig of 1411 bp in length 1412 1511: gap of 100 bp 1512 2165: contig of 654 bp in length 2166 2265: gap of 100 bp 2266 3074: contig of 809 bp in length 3075 3174: gap of 100 bp 3175 4763: contig of 1589 bp in length 4764 4863: gap of 100 bp 4864 6161: contig of 1298 bp in length 6162 6261: gap of 100 bp

*	6252	8601:	contig of 2340 bp	in length
*	8602	8701:	gap of 100 bp	
*	8702	13139:	contig of 4438 bp	in length
*	13140	13239:	gap of 100 bp	
*	13240	17613:	contig of 4374 bp	in length
*	17614	17713:	gap of 100 bp	
*	17714	30894:	contig of 13181 bp	in length
*	30895	30994:	gap of 100 bp	
*	30995	40051:	contig of 9057 bp	in length
*	40052	40151:	gap of 100 bp	
*	40152	95859:	contig of 53708 bp	in length
*	95860	95959:	gap of 100 bp	
*	95960	138459:	contig of 42500 bp	in length

FEATURES

```

source
1. .i38459
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-254K5"
/clone_lib="RPC1-11 Human Male BAC"
1. .i1411
misc_feature
  clone="assembly_fragment"
  clone_end:SP6
  vector_side:left"
  1512..2165
  /note="assembly_fragment"
  -2266..3074
  /note="assembly_fragment"
  3175..4763
  /note="assembly_fragment"
  4864..6161
  /note="assembly_fragment"
  6262..8601
  /note="assembly_fragment"
  8702..13139
  /note="assembly_fragment"
  13240..17613
  /note="assembly_fragment"
  17714..30894
  /note="assembly_fragment"
  30995..40051
  /note="assembly_fragment"
  40152..95859
  /note="assembly_fragment"
  95960..138459
  /note="assembly_fragment"
  clone_end:r7
  vector_side:right"
BASE COUNT      38140 a 31260 c 40041 g 37914 t 1104 others

```

Query Match	49.2%;	Score 492;	DB 2;	Length 138459;
Best Local Similarity	100.0%;	Pred. No. 5e-280;		
Matches 492;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
<hr/>				
QY	509	AAATATAAATAAAAATAAAAAATAAAATAAAATAAAATAAAATCTTAAAGGGG	568	
Db	34048	AAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATCTTAAAGGG	34107	
<hr/>				
QY	569	CTGGCAATTGGCTAGCACCTTATATGCCCAATAGTAATAGCTATCAATATCCCCACCCCT	628	
<hr/>				
Db	34108	CTGGCAATTGGCCTAGCACCTTATATGCCCAATAGTAATAGCTATCAATATCCCCACCCCT	34167	
<hr/>				
QY	629	ACCACGTGCTGAATTTAGTTTTCTTTTTTGACCCCCCATTAGACTTAAGCGAATTC	688	
<hr/>				
Db	34168	ACCACGTGCTGAATTTAGTTTTCTTTTTTGACCCCCCATTAGACTTAAGCGAATTC	34227	
<hr/>				
QY	689	TCACCGTACTCCTCTGTAATTTCTGGTCTCTGGCACATAGTTGGGTCTCAGTGAACAT	748	
<hr/>				
Db	34228	TCACCGTACTCCTCTGTAATTTCTGGTCTCTGGCACATAGTTGGGTCTCAGTGAACAT	34287	
<hr/>				
QY	749	GGTGAGTGAATGCACAATGCAAGGAATCTCCAGGCATCTGGGAGCCCTCCAGGCGGG	808	

Db	34288	GTGAGTGAATGAGCAATTCAGGAAATCTCCAGGCATCTGGAGGCCCTCCCGAGGGG	34347
Qy	809	TGAGTTCGGGAAACTCATAGTCGTCTCAATGGCCCACTGAAAGGTAGAGAGTTCTCTGGG	868
Db	34348	TGAGTTCGGGAAACTCATAGTCGTCTCAATGGCCCACTGAAAGGTAGAGAGTTCTCTGGG	34407
Qy	869	TCCCACTTCGGCACCCCATCTCTGACTCACTGCTGAAATAATAATAATAATAAATAAAT	928
Db	34408	TCCCACTTCGGCACCCCATCTCTGACTCACTGCTGAAATAATAATAATAATAAATAAAT	34467
Qy	929	ACACTTATCCGGAGCGCTCCCAATGCCCTTCGCAGGATGCGACAGGCCACGCAATGA	988
Db	34468	ACACTTATCCGGAGCGCTCCCAATGCCCTTCGCAGGATGCGAAGGAGCCGACGAATGA	34527
Qy	989	TGACCGCGCTGC	1000
Db	34528	TGACCGCGCTGC	34539

RESULT 4
AC104363

LOCUS	AC104363	62102 bp	DNA	linear	HTG 08-DEC-2001			
DEFINITION	Homo sapiens chromosome 18 clone RP11-42N11 map 18, LOW-PASS SEQUENCE SAMPLING.							
ACCESSION	AC104363							
VERSION	AC104363.1	GI:117426352						
KEYWORDS	HTG; HGVS-PHASE0.							
SOURCE	Homo sapiens.							
ORGANISM	Homo sapiens							
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.							
REFERENCE	1 (bases 1 to 62102)							
AUTHORS	Birren,B., Linton,L., Nusbaum,C. and Lander,E.							
TITLE	Homo sapiens chromosome 18, clone RP11-42N11							

REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
REFERENCE	
AUTHORS	

Blurred, B., Lincoln, L., Nussbaum, C., Lander, E., All, A., Allen, N.,
 Anderson, S., Barina, N., Bastien, V., Boguslavsky, L., Bouhgalter, B.,
 Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
 Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A.,
 Cooke, P., DeAtellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
 Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hago, B., Heard, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
 Jones, C., Katat, A., Karatas, A., Kells, C., LaRocque, K.,
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
 MacLean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J.,
 Meneus, L., Mihova, T., Mlenga, T., Murphy, T., Naylor, J., Nguyen, C.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,
 Seaman, S., Severy, P., Sencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
 Zainoun, J., Zemlek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (08-DEC-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RW/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WTBFR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L22094
Center clone name: 42 N 11

 * NOTE: This record contains 81 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

```

1      665: contig of 665 bp in length
*      666 765: gap of 100 bp
*      766      1454: contig of 689 bp in length
*      1455 1554: gap of 100 bp
*      1555      2232: contig of 678 bp in length
*      2233 2332: gap of 100 bp
*      2333      3028: contig of 696 bp in length
*      3029 3128: gap of 100 bp
*      3129 3753: contig of 625 bp in length
*      3754 3853: gap of 100 bp
*      3854      4459: contig of 606 bp in length
*      4460 4559: gap of 100 bp
*      4560      5225: contig of 666 bp in length
*      5226 5325: gap of 100 bp
*      5326 5979: contig of 654 bp in length
*      5980 6079: gap of 100 bp
*      6080      6752: contig of 673 bp in length
*      6753 6852: gap of 100 bp
*      6853      7526: contig of 674 bp in length
*      7527 7626: gap of 100 bp
*      7627      8307: contig of 681 bp in length
*      8308 8407: gap of 100 bp
*      8408      9090: contig of 683 bp in length
*      9091 9190: gap of 100 bp
*      9191      9880: contig of 690 bp in length
*      9881 9980: gap of 100 bp
*      9981      10667: contig of 687 bp in length
*      10668 10767: gap of 100 bp
*      10768 11436: contig of 669 bp in length
*      11437 11536: gap of 100 bp
*      11537 12157: contig of 621 bp in length
*      12158 12257: gap of 100 bp
*      12258 12913: contig of 656 bp in length
*      12914 13013: gap of 100 bp
*      13014 13667: contig of 654 bp in length
*      13668 13767: gap of 100 bp
*      13768 14429: contig of 662 bp in length
*      14430 14529: gap of 100 bp
*      14530 15189: contig of 660 bp in length
*      15190 15289: gap of 100 bp
*      15290 15950: contig of 661 bp in length
*      15951 16050: gap of 100 bp
*      16051 16724: contig of 674 bp in length
*      16725 16824: gap of 100 bp
*      16825 17505: contig of 681 bp in length
*      17506 17605: gap of 100 bp
*      17606 18290: contig of 685 bp in length
*      18291 18390: gap of 100 bp
*      18391 19050: contig of 660 bp in length
*      19051 19150: gap of 100 bp
*      19151 19807: contig of 657 bp in length
*      19808 19907: gap of 100 bp
*      19908 20568: contig of 661 bp in length
*      20569 20668: gap of 100 bp
*      20669 21322: contig of 654 bp in length
*      21323 21422: gap of 100 bp
*      21423 22079: contig of 657 bp in length
*      22080 22179: gap of 100 bp
*      22180 22852: contig of 673 bp in length
*      22853 22952: gap of 100 bp
*      22953      23621: contig of 669 bp in length

```

```

23622 23721: gap of 100 bp
*      23722      24386: contig of 665 bp in length
*      24387 24486: gap of 100 bp
*      24487      25174: contig of 688 bp in length
*      25175 25274: gap of 100 bp
*      25275      25957: contig of 683 bp in length
*      25958 26057: gap of 100 bp
*      26058      26729: contig of 672 bp in length
*      26730 26829: gap of 100 bp
*      26830      27478: contig of 649 bp in length
*      27479 27578: gap of 100 bp
*      27579      28237: contig of 659 bp in length
*      28238 28337: gap of 100 bp
*      28338      29002: contig of 665 bp in length
*      29003 29102: gap of 100 bp
*      29103      29764: contig of 662 bp in length
*      29765 29864: gap of 100 bp
*      29865      30544: contig of 680 bp in length
*      30545 30644: gap of 100 bp
*      30645      31330: contig of 686 bp in length
*      31331 31430: gap of 100 bp
*      31431      32107: contig of 677 bp in length
*      32108 32207: gap of 100 bp
*      32208      32887: contig of 680 bp in length
*      32888 32987: gap of 100 bp
*      32988      33679: contig of 692 bp in length
*      33680 33779: gap of 100 bp
*      33780      34451: contig of 672 bp in length
*      34452 34551: gap of 100 bp
*      34552      35216: contig of 665 bp in length
*      35217 35316: gap of 100 bp
*      35317      35978: contig of 662 bp in length
*      35979 36078: gap of 100 bp
*      36079      36731: contig of 653 bp in length
*      36732 36831: gap of 100 bp
*      36832      37486: contig of 655 bp in length
*      37487 37586: gap of 100 bp
*      37587      38249: contig of 663 bp in length
*      38250 38349: gap of 100 bp
*      38350      39032: contig of 683 bp in length
*      39033 39132: gap of 100 bp
*      39133      39815: contig of 683 bp in length
*      39816 39915: gap of 100 bp
*      39916      40576: contig of 661 bp in length
*      40577 40676: gap of 100 bp
*      40677      41356: contig of 680 bp in length
*      41357 41456: gap of 100 bp
*      41457      42141: contig of 685 bp in length
*      42142 42241: gap of 100 bp
*      42242      42899: contig of 658 bp in length
*      42900 42999: gap of 100 bp
*      43000      43650: contig of 651 bp in length
*      43651 43750: gap of 100 bp
*      43751      44412: contig of 662 bp in length
*      44413 44512: gap of 100 bp
*      44513      45169: contig of 657 bp in length
*      45170 45269: gap of 100 bp
*      45270      45925: contig of 656 bp in length
*      45926 46025: gap of 100 bp
*      46026      46695: contig of 670 bp in length
*      46696 46795: gap of 100 bp
*      46796      47481: contig of 686 bp in length
*      47482 47581: gap of 100 bp
*      47582      48185: contig of 604 bp in length
*      48186 48285: gap of 100 bp
*      48286      48964: contig of 679 bp in length
*      48965 49064: gap of 100 bp
*      49065      49751: contig of 687 bp in length
*      49752 49851: gap of 100 bp
*      49852      50518: contig of 667 bp in length
*      50519 50618: gap of 100 bp
*      50619      51290: contig of 672 bp in length
*      51291 51390: gap of 100 bp

```



```
repeat_region 17712..17724
/note="FAM_C repeat: matches 130..142 of consensus"
18035..18164
/note="FAM_C repeat: matches 1..130 of consensus"
19499..19566
/note="MUT1A2-internal repeat: matches 1582..1651 of
consensus"
repeat_region 19567..19870
/note="AluSc repeat: matches 1..294 of consensus"
20838..20934
/note="L1M1 repeat: matches -1210..-1134 of consensus"
join(24239..24291,29142..29175,29299..29496,30784..30942,
33601..33681,35749..35805,41891..41933,42579..42769,
43394..43715,44666..44724,46107..46206,46993..47154,
49581..49786,51645..51745,52698..52879,55139..55224,
59047..59180,59335..59440,59532..59640,59921..60054,
60223..60386,60755..61007)
/gene="dj979N1.1"
/codon_start=1
/evidence=not_experimental
/protein_id="dj979N1.1 (dj979N1.1)"
/db_xref="GI:6562060"
/translation="MERQKADIEKGLQFQSTLPKQVEYEFLLKIVNLFASGN
DLFREKYQALVYMEGLNVAADYASDOVALPRELLCKLRVNRACVFMGLYERKAL
DESEKALGDSIESIRALFRKARALNELGRHKEAYECSSRCSALPHDESVTQGEQA
OKLGRVKRKYRQPELETFSLLSNGTAGVADQGTSGSLGSDIDTETGNVPTREQV
EIGAPDCYVDRGSPALLPSTPMPHPVLDLLAPLDSRTLSLDSLDSFSDGV
FGELDTLLDSLSVQGLSGSGVSELPQIPVFRGTPPLPPVVGSGSPVSPPPP
ASFLGVMDPSKLAASVLDALDPPGTDLPDLPLPYSETRLDALDSGSTRGSLKDP
SFMETNSQDHRPPSGAQPAPSPPCMPNTALLIKNPLAATHEKQACQICPKTGP
RAGDITYREGLEHKRDILLGLRSEDOQWRRIRPRETKYSFVGSYILCKDMINKQ
DKYGDNCTEAYHOEIDVWBERGTLNROLLDFLGGVKKSGTIAKLLKHOGIF
TFELCEICDSPRIISGNDKSPVCSNLAHKSFYNNKCLVHIVRSTLSKYSKIRQF
QEFQFDVREHYRGCLRDESCFAHFSFELKVLWLLQYQSGMTHDIVESKYYWQ
MEAHGAKSSMGAPRTHGFTFDLQMKFVCGQWNGOVVPDKOLKYSARHWO
TERRVLLVMSKARKWSVRPLIRNFQOYDLICHAQNGRCKQYVGNCSFASPE
ERDMTFMKNKILDMQOYDMLAKHNPCKPGEGTPISSREGEKQIOMPTDYADIMG
GYHCLGCKNSKKQOQHIOSEKHEKFTSDSDASGAFRPFMGFEFLCDLQMG
KACPGDKCRCAHGQBELEWLDRLVKLQKAKARKMLLCPDRDDDFGRYNFLLED
GDAAGATPEAPAAAATTTGE"
23406..23716
/note="AluYa5 repeat: matches 1..311 of consensus"
26153..26298
/note="MER69A repeat: matches 1..133 of consensus"
28743..29117
/note="L1M3e repeat: matches -429..737 of consensus"
complement(30252..30637)
/note="match: STS: Em:G14621"
join(33626..33681,35749..35805,41891..41933,42579..42630)
/gene="dj979N1.2"
join(<33626..33681,35749..35805,41891..41933,
42579..42630)
/gene="dj979N1.2"
/note="supported by GENSCAN
GENSCAN, CpG island and polyA features suggest this gene
fragment and fragments dj979N1.1, -.3 and -.4 are part of
one gene
match: ESTs: Em:AI122015 Em:AA306154"
/codon_start=3
/evidence=not_experimental
/product="dj979N1.2 (novel protein)"
/protein_id="CAB62494.1"
/db_xref="GI:6562061"
/translation="ELAQKGLRVRKAYRKPQETFSLLSNGTAGVADQGTNSGLG
SIDDIETDCYVDPRGSPALLPSTPT"
34507..35006
/gene="dj979N1.2"
/note="CpG island"
/evidence=not_experimental
34636..34823
/gene="dj979N1.2"
/note="Single clone region. region contains only reads

lm07 which is a SIL from subclone 22zs76e5, assembly
consistent with restriction digest."
join(46993..47154,49581..49786,51645..51745,52698..52879,
55139..55224)
/gene="dj979N1.3"
join(<46993..47154,49581..49786,51645..51745,52698,
55139..55224)
/gene="dj979N1.3"
/note="supported by GENES and GENSCAN
GENSCAN (and partially FGENES), CpG island and polyA
features suggest that this gene fragment and fragments
dj979N1.1, -.2 and -.4 are part of one gene
match: ESTs: Em:R88846 Em:AA624869 Em:W78795 Em:AI138030
Em:AI642278 Em:H55396 Em:AI791927 Em:H55347 Em:H55369
Em:AI821160"
/codon_start=3

Query Match 7.7%; Score 77; DB 9; Length 102151;
Best Local Similarity 100.0%; Pred. No. 1.le-33;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 335 GTCCTCTACTAAAAATACAAAATATAGCCAGGTGTGTGGCACACGCTGTAGTCCCGACT 394
|||||
Db 18376 GTCCTCTACTAAAAATACAAAATATAGCCAGGTGTGTGGCACACGCTGTAGTCCCGACT 18317
|||||
QY 395 ACTTGGGAGCGTGAGGC 411
|||||
Db 18316 ACTTGGGAGCGTGAGGC 18300
|||||

RESULT 6
AC016395/c
AC016395
LOCUS 166869 bp DNA linear PRI 14-MAY-2002
DEFINITION Homo sapiens chromosome 10 clone RP11-153K11, complete sequence.
ACCESSION AC016395
VERSION AC016395.5 GI:20564406
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Smith,D.R.
Genome Therapeutics Corporation Sequencing Center: Human Genome
Sequence Data
Unpublished
2 (bases 1 to 166869)
Smith,D.R.
Direct Submission
Submitted (26-NOV-1999) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
3 (bases 1 to 166869)
Smith,D.R.
Direct Submission
Submitted (14-MAY-2002) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
On May 14, 2002 this sequence version replaced gi:9929646.
Location/Qualifiers
1..166869
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-153K11"
/clone_lib="RPC1-11"
BASE COUNT 47408 a 35216 c 35336 g 48909 t
ORIGIN

Query Match 7.6%; Score 76; DB 9; Length 166869;
```

```

Best Local Similarity 100.0%; Pred. No. 4.5e-33;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 336 TCTCTACTAAATATACAAAATAGCCAGGTGTGGTGGCACACGCTGTAGTCCAGCTA 395
|||||
Db 73964 TCTCTACTAAATATACAAAATAGCCAGGTGTGGTGGCACACGCTGTAGTCCAGCTA 73905
|||||
QY 396 CTCTGGAGGCTGAGGC 411
|||||
Db 73904 CTCTGGAGGCTGAGGC 73889

RESULT 7
AL3900071
LOCUS      AL3900071               170998 bp    DNA    linear    PRI 25-MAR-2001
DEFINITION Human DNA sequence from clone RP11-98D3 on chromosome 13, complete
sequence.
ACCESSION  AL3900071
VERSION    AL3900071.9  GI:13373975
KEYWORDS   HTG.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 170998)
            Phillimore, B.
            Direct Submission
            Submitted (23-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
            On Mar 16, 2001 this sequence version replaced gi:12743811.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TrEMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep
            This sequence was generated from part of bacterial clone contigs of human
            chromosome 13, constructed by the Sanger Centre Chromosome 13
            Mapping Group. Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr13
            RP11-98D3 is from the library RP11-98D3 constructed by the group of
            Pieter de Jong. For further details see
            http://www.choi.org/bacpac/home.htm
            VECTOR: pBACE3.6
            This sequence is the entire insert of clone RP11-98D3 The true left
            end of clone RP11-88019 is at 112608 in this sequence. The true
            right end of clone RP11-29G24 is at 33813 in this sequence.
FEATURES
            source
                1..170998
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /chromosome="13"
                    /clone="RP11-98D3"
                    /clone_lib="RP11-11"
                    /vector="pBACE3.6"
            misc_feature
                1..132
                    /note="match: STS: Em:HSPH27EB"
            repeat_region
                1332..1494
                    /note="MER20 repeat: matches 73..217 of consensus"
            repeat_region
                1495..1736
                    /note="L2 repeat: matches 2339..2594 of consensus"

```

```

repeat_region
1848..1965
/note="L1P45 repeat: matches 6026..6143 of consensus"
2110..2412
/note="AluSq repeat: matches 1..303 of consensus"
4254..4301
/note="L2 copies 4 mer ttc 75% conserved"
7203..7239
/note="L2 repeat: matches 2620..2656 of consensus"
7563..7708
/note="MIR repeat: matches 110..257 of consensus"
7989..8060
/note="3 copies 24 mer 80% conserved"
8003..8060
/note="29 copies 2 mer at 91% conserved"
8005..8060
/note="14 copies 4 mer atat 92% conserved"
8062..8718
/note="L1P45 repeat: matches 5486..6142 of consensus"
9019..9070
/note="13 copies 4 mer tata 80% conserved"
10761..10991
/note="L2 repeat: matches 2482..2710 of consensus"
11163..11609
/note="L2 repeat: matches 2010..2492 of consensus"
12426..12736
/note="AluJb repeat: matches 1..299 of consensus"
13347..13654
/note="AluY repeat: matches 1..309 of consensus"
14005..14129
/note="MIR repeat: matches 19..148 of consensus"
16812..17120
/note="AluY repeat: matches 1..309 of consensus"
17927..18055
/note="MIR repeat: matches 86..218 of consensus"
19169..19260
/note="46 copies 2 mer at 69% conserved"
19195..19254
/note="15 copies 4 mer atat 76% conserved"
19496..19845
/note="L1MB8 repeat: matches 5830..6173 of consensus"
19846..20139
/note="AluY repeat: matches 1..295 of consensus"
20140..20489
/note="L1MB8 repeat: matches 5462..5830 of consensus"
23000..23410
/note="MSTRC repeat: matches 15..405 of consensus"
24771..25102
/note="L2 repeat: matches 1900..2247 of consensus"
25229..25415
/note="L2 repeat: matches 1485..1683 of consensus"
25574..25874
/note="AluJb repeat: matches 1..304 of consensus"
26566..26693
/note="L2 repeat: matches 2352..2476 of consensus"
26816..26994
/note="MIR repeat: matches 29..226 of consensus"
27529..27694
/note="L2 repeat: matches 2495..2644 of consensus"
28745..29001
/note="AluJo repeat: matches 1..256 of consensus"
29425..29716
/note="AluSq repeat: matches 1..292 of consensus"
30387..30983
/note="MER90 repeat: matches 1..615 of consensus"
31050..31356
/note="AluX repeat: matches 1..306 of consensus"
32941..33577
/note="CpG island"
/evidence=not_experimental
37667..38255
/note="CpG island"
/evidence=not_experimental
38473..38928

```

```

repeat_region /note="match: STS: Em:G09543"
38530. 38589
/misc_feature /note="20 copies 3 mer ctg 98% conserved"
40650. .41215
/note="CpG island"
repeat_region /evidence=not_experimental
44973. .45123
/note="L1M1 repeat: matches 6016. .6159 of consensus"
46468. .46527
/note="30 copies 2 mer tg 85% conserved"
48732. .48843
/note="MIR repeat: matches 60. .196 of consensus"
52923. .53162
/note="MIR repeat: matches 11. .245 of consensus"
53167. .53264
/note="MIR repeat: matches 106. .213 of consensus"
53297. .53386
/note="MIR repeat: matches 8. .98 of consensus"
53543. .53889
/note="L2 repeat: matches 1651. .2047 of consensus"
54290. .54575
/note="L1P81 repeat: matches 5852. .6155 of consensus"
56551. .56667
/note="L1M4 repeat: matches 2267. .2282 of consensus"
56668. .56746
/note="MAD81 repeat: matches 1. .80 of consensus"
56747. .56968
/note="L1M4 repeat: matches 2282. .2155 of consensus"
56969. .57270
/note="AluYb repeat: matches 4. .311 of consensus"
57271. .57702
/note="L1M4 repeat: matches 2154. .2583 of consensus"
57788. .57860
/note="MER47 repeat: matches 1. .73 of consensus"
57862. .58057
/note="L1M4 repeat: matches 2600. .2811 of consensus"
58394. .58520
/note="FLAM_A repeat: matches 1. .121 of consensus"
58586. .58893
/note="AluYb repeat: matches 1. .311 of consensus"
64210. .64446
/note="L2 repeat: matches 2453. .2706 of consensus"
64871. .64953
/note="MER63 repeat: matches 685. .766 of consensus"
66867. .67144
/note="AluYb repeat: matches 1. .299 of consensus"
68291. .68591
/note="AluYb repeat: matches 11. .305 of consensus"
69232. .69973
/note="L2 repeat: matches 1244. .2019 of consensus"
70272. .70418
/note="MIR repeat: matches 23. .172 of consensus"
70713. .71105
/note="L2 repeat: matches 1629. .2037 of consensus"
71119. .71618
/note="MER54B repeat: matches 1. .498 of consensus"
71602. .71825
/note="MER54B repeat: matches 459. .674 of consensus"
71894. .72125
/note="L2 repeat: matches 1300. .1544 of consensus"
72923. .73109
/note="MIR repeat: matches 58. .249 of consensus"
75367. .75683
/note="AluX repeat: matches 1. .303 of consensus"
78989. .79225
/note="MIR repeat: matches 12. .256 of consensus"
80308. .80410
/note="L2 repeat: matches 2577. .2692 of consensus"

```

Query Match 7.6%; Score 76; DB 9; Length 170998;
 Best Local Similarity 100.0%; Pred.No. 4.5e-33;
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 336 TCTCTACTAAATAACAAAATAGCCAGGTGTGGCAGCACCGCTAGTCCAGCTA 395
|||||
Db 163825 TCTCTACTAAATAACAAAATAGCCAGGTGTGGCAGCACCGCTAGTCCAGCTA 163884
|||||

QY 396 CTTGGAGGCTGAGGC 411
|||||
Db 163885 CTTGGAGGCTGAGGC 163900
|||||

RESULT 8
AL442068
LOCUS Homo sapiens chromosome 10 clone RP11-331G8, *** SEQUENCING IN
DEFINITION PROGRESS ***, 43 unordered pieces.
ACCESSION AL442068
VERSION AL442068.16 GI:17043795
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Sims.S.
Direct Submission
Submitted (19-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Nov 21, 2001 this sequence version replaced gi:14575284.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA331G8
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: dye-terminator Big Dye; 100% of reads
Consensus quality: 277809 bases at least Q40
Consensus quality: 281712 bases at least Q30
Consensus quality: 284299 bases at least Q20
Insert size: 286425; sum-of-contigs
Insert size: 156064; 7.4% error; agarose-fp
Quality coverage: 5.79x in Q20 bases; sum-of-contigs Quality
coverage: 15.41x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 43 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 103320: contig of 103320 bp in length
* 103321 103420: gap of 100 bp
* 103421 140724: contig of 37304 bp in length
* 140725 140824: gap of 100 bp
* 140825 162687: contig of 21863 bp in length
* 162688 162787: gap of 100 bp
* 162788 165269: contig of 2482 bp in length
* 165270 165369: gap of 100 bp
* 165370 167787: contig of 2418 bp in length
* 167788 167887: gap of 100 bp
* 167888 170672: contig of 2785 bp in length
* 170673 170772: gap of 100 bp
* 170773 173340: contig of 2568 bp in length
* 173341 173440: gap of 100 bp
* 173441 176369: contig of 2929 bp in length
* 176370 176469: gap of 100 bp
* 176470 179660: contig of 3191 bp in length
* 179661 179760: gap of 100 bp

```

```

* 179761 181788: contig of 2028 bp in length
* 181789 181888: gap of 100 bp
* 181889 185591: contig of 3703 bp in length
* 185592 185691: gap of 100 bp
* 185692 187897: contig of 2206 bp in length
* 187898 187997: gap of 100 bp
* 187998 190151: contig of 2154 bp in length
* 190152 190251: gap of 100 bp
* 190252 192977: contig of 2726 bp in length
* 192978 193077: gap of 100 bp
* 193078 195504: contig of 2427 bp in length
* 195505 198738: contig of 3134 bp in length
* 198739 198838: gap of 100 bp
* 198839 202505: contig of 3667 bp in length
* 202506 202605: gap of 100 bp
* 202606 206757: contig of 4152 bp in length
* 206758 206857: gap of 100 bp
* 206858 211910: contig of 5053 bp in length
* 211911 212010: gap of 100 bp
* 212011 218152: contig of 6142 bp in length
* 218153 218252: gap of 100 bp
* 218253 221882: contig of 3630 bp in length
* 221883 221982: gap of 100 bp
* 221983 225735: contig of 3753 bp in length
* 225736 225835: gap of 100 bp
* 225836 229907: contig of 4072 bp in length
* 229908 230007: gap of 100 bp
* 230008 233016: contig of 3009 bp in length
* 233017 233116: gap of 100 bp
* 233117 236145: contig of 3029 bp in length
* 236146 236245: gap of 100 bp
* 236246 239609: contig of 3364 bp in length
* 239610 239709: gap of 100 bp
* 239710 241796: contig of 2087 bp in length
* 241797 241896: gap of 100 bp
* 241897 243915: contig of 2019 bp in length
* 243916 244015: gap of 100 bp
* 244016 246442: contig of 2427 bp in length
* 246443 246542: gap of 100 bp
* 246543 249599: contig of 3057 bp in length
* 249600 249699: gap of 100 bp
* 249700 252680: contig of 2981 bp in length
* 252681 252780: gap of 100 bp
* 252781 256312: contig of 3532 bp in length
* 256313 256412: gap of 100 bp
* 256413 259563: contig of 3151 bp in length
* 259564 259663: gap of 100 bp
* 259664 262204: contig of 2541 bp in length
* 262205 262304: gap of 100 bp
* 262305 264799: contig of 2495 bp in length
* 264800 264899: gap of 100 bp
* 264900 267645: contig of 2746 bp in length
* 267646 267745: gap of 100 bp
* 267746 271525: contig of 3780 bp in length
* 271526 271625: gap of 100 bp
* 271626 275798: contig of 4173 bp in length
* 275799 275898: gap of 100 bp
* 275899 278025: contig of 2127 bp in length
* 278026 278125: gap of 100 bp
* 278126 281433: contig of 3308 bp in length
* 281434 281533: gap of 100 bp
* 281534 284437: contig of 2904 bp in length
* 284438 284537: gap of 100 bp
* 284538 286727: contig of 2190 bp in length
* 286728 286827: gap of 100 bp
* 286828 290625: contig of 3798 bp in length.
Location/Qualifiers
1..230625
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-331G8"

```

FEATURES

source

```

misc_feature
1..103320
/clone_lib="RPC1-11.2"
/note="assembly_fragment:07950"
fragment_chain:1
vector_side:left"
103421..140724
/note="assembly_fragment:01658"
fragment_chain:1"
140825..162687
/note="assembly_fragment:01866"
fragment_chain:1"
162788..165269
/note="assembly_fragment:01933"
fragment_chain:2"
165370..167787
/note="assembly_fragment:02386"
fragment_chain:2"
167888..170672
/note="assembly_fragment:02267"
fragment_chain:3"
170773..173340
/note="assembly_fragment:02094"
fragment_chain:3"
173441..176369
/note="assembly_fragment:00062"
176470..179660
/note="assembly_fragment:00157"
179761..181788
/note="assembly_fragment:00471"
181889..185591
/note="assembly_fragment:00485"
185692..187897
/note="assembly_fragment:00607"
187998..190151
/note="assembly_fragment:01019"
190252..192977
/note="assembly_fragment:01432"
193078..195504
/note="assembly_fragment:01879"
195605..198738
/note="assembly_fragment:01944"
198839..202505
/note="assembly_fragment:02174"
202606..206757
/note="assembly_fragment:02200"
206858..211910
/note="assembly_fragment:02684"
212011..218152
/note="assembly_fragment:03379"
218253..221882
/note="assembly_fragment:03527"
221983..225735
/note="assembly_fragment:03880"
225836..229907
/note="assembly_fragment:04098"
230008..233016
/note="assembly_fragment:04166"
233117..236145
/note="assembly_fragment:04570"
236246..239609
/note="assembly_fragment:04709"
239710..241796
/note="assembly_fragment:04939"
241897..243915

```

Query Match
 Best Local Similarity 100.0%; DB 2: Length 290625;
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 336 TCTCTACTAAAAATACAAAAATTAGCCAGGTGTGGTGGCAGCAGCCTGTAGTCCAGCTA 395
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 254811 TCTCTACTAAAAATACAAAAATTAGCCAGGTGTGGTGGCAGCAGCCTGTAGTCCAGCTA 254870

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

RESULT 11

AC011505
 LOCUS 167417 bp DNA linear PRI 14-JUL-2002
 DEFINITION Homo sapiens chromosome 19 clone CTD-2081K17, complete sequence.
 AC011505
 VERSION AC011505.6 GI:21747445
 KEYWORDS HTG.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 DOE Joint Genome Institute and Stanford Human Genome Center.
 Direct Submission
 Unpublished
 2 (bases 1 to 167417)
 DOE Joint Genome Institute.
 Direct Submission
 Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 3 (bases 1 to 167417)
 DOE Joint Genome Institute and Stanford Human Genome Center.
 Direct Submission
 Submitted (01-JUN-2002) DOE Joint Genome Institute, 2800 Mitchell
 Drive, Walnut Creek, CA 94598, USA
 4 (bases 1 to 167417)
 DOE Joint Genome Institute and Stanford Human Genome Center.
 Direct Submission
 Submitted (14-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell
 Drive, Walnut Creek, CA 94598, USA
 On Jul 14, 2002 this sequence version replaced gi:21306640.
 Draft Sequence Produced by DOE Joint Genome Institute
 www.jgi.doe.gov
 Finishing Completed at Stanford Human Genome Center
 www.shgc.stanford.edu
 Quality: Phrap Quality >=40 99.9% of Sequence;
 Estimated Total Number of Errors is 0.1.
 NOTE: Shatter libraries failed to resolve the dinucleotide repeat
 region. Unsure number of repeat copies 49544-50135. False join
 49558.

FEATURES

Location/Qualifiers

1. .167417
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="19"
 /clone="CTD-2081K17"
 49544..50135
 /note="NOTE: Shatter libraries failed to resolve the
 dinucleotide repeat region. Unsure number of repeat
 copies 49544-50135. False join 49658."
 49576 a 40334 c 40235 g 41872 t

misc_feature

BASE COUNT 44976 a 40334 c 40235 g 41872 t
 ORIGIN
 Query Match 7.4%; Score 74; DB 9; Length 167417;
 Best Local Similarity 100.0%; Pred. No. 6.9e-32;
 Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 335 GTCTCTACTAAATAACAAATTAGCCAGGTGTGGTGGCACACGCTGTAGTCCAGCT 394
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 36358 GTCTCTACTAAATAACAAATTAGCCAGGTGTGGTGGCACACGCTGTAGTCCAGCT 36417

QY 395 ACTTGGGAGGCTGA 408
 ||||||||||||

Db 36418 ACTTGGGAGGCTGA 36431

RESULT 12

AF190465
 LOCUS 96308 bp DNA linear PRI 27-OCT-1999
 DEFINITION Homo sapiens SH3-containing protein EEN (EEN) and chromatin
 assembly factor-I p150 subunit (CAF) genes, complete cds, complete
 sequence.
 AF190465

ACCESSION

AF190465.1 GI:6120105
 KEYWORDS HTG.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 96308)
 Chen, S., Xiong, H., Dong, H., Lin, W., Zhang, C., Fu, G., Qi, Z. and
 Huang, G.M.
 Homo sapiens SH3-containing protein EEN gene
 Unpublished
 2 (bases 1 to 96308)
 Chen, S., Xiong, H., Dong, H., Lin, W., Zhang, C., Fu, G., Qi, Z. and
 Huang, G.M.
 Direct Submission
 Submitted (28-SEP-1999) Chinese Human Genome Center at Shanghai,
 351 Guoshoujing Road, Zhangjiang Hi-Tech Park, Shanghai 201203,
 People's Republic of China
 Location/Qualifiers
 1. .96308
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="19"
 /map="19p13"
 /clone="BAC 335C9"
 complement(<2348..>.41096)
 /gene="EEN"
 gene
 complement(join(<2348..2544,3075..3131,3358..3482,
 4116..4219,4464..4622,4832..4965,6230..6373,7250..7322,
 7675..7743,41052..>.41096))
 /gene="EEN"
 /product="SH3-containing protein EEN"
 /complement(join(2348..2544,3075..3131,3358..3482,
 4116..4219,4464..4622,4832..4965,6230..6373,7250..7322,
 7675..7743,41052..41096))
 /gene="EEN"
 /codon_start=1
 /product="SH3-containing protein EEN"
 /protein_id="AAF04290.1"
 /db_xref="GI:6120106"
 /translation="MSVAGLKKQFYKASOLYSEKVGAGETKLDODFKEMKKVDVTS
 KAVTEVLARTIEYLOPNPASRAKLWLNVTSKIQGVKNPGYDSEGLGECMRHGK
 ELGGSENFGLDLDAGSMKELAYVDSLDIEVKQNFIDPLQNLCEKDKETQHLKK
 LEGRRLDFDYKKKGKQIPDELRQALEKFEESKEVAETSMNLLETIEQVSLGAL
 VDAQLDYHRQAVQILDELAELKLRMRREASSRKREYKPKPREPFDLGFQSNNGFP
 CTTAPKAASSFRSSDKPIRTPSRSMPLDQPSCKALYDFEPENDGELGFHEDGVT
 LTNQIDENWYEGMLDGGSGFFPLSYVEVLVLPQ"
 complement(<2348..2544)
 /gene="EEN"
 /number=10
 complement(3075..3131)
 /gene="EEN"
 /number=9
 complement(3358..3482)
 /gene="EEN"
 /number=8
 complement(4116..4219)
 /gene="EEN"
 /number=7
 complement(4464..4622)
 /gene="EEN"
 /number=6
 complement(4832..4965)
 /gene="EEN"
 /number=5
 complement(5036..5177)
 /rpt_family="FRAM"
 complement(5190..5471)
 /rpt_family="Alusg"
 5473..5501
 /rpt_family="(TTTTG)n"
 complement(5509..5611)
 /rpt_family="Alusp"
 repeat_region
 complement(5036..5177)
 /rpt_family="FRAM"
 complement(5190..5471)
 /rpt_family="Alusg"
 5473..5501
 /rpt_family="(TTTTG)n"
 complement(5509..5611)
 /rpt_family="Alusp"

```

repeat_region 5612...5654
                /rpt_family="TG)n"
repeat_region complement(5655..5803)
                /rpt_family="AluSp"
exon           complement(6230..6373)
                /gene="EEN"
exon           /number=4
                complement(7250..7322)
                /gene="EEN"
exon           /number=3
                complement(7675..7743)
                /gene="EEN"
repeat_region /number=2
                /rpt_family="MER5B"
repeat_region 9484..9533
                /rpt_family="MER5B"
repeat_region 9540..9847
                /rpt_family="AluY"
repeat_region complement(12727..12855)
                /rpt_family="FLAM_C"
repeat_region 13851..13882
                /rpt_family="FLAM_C"
repeat_region 18847..18905
                /rpt_family="AGGGGG)n"
repeat_region /rpt_family="L2"
                /rpt_family="L2"
repeat_region 19057..19141
                /rpt_family="MIR"
repeat_region 19180..19487
                /rpt_family="AluSp"
repeat_region 19921..20192
                /rpt_family="AluSx"
repeat_region 21780..22103
                /rpt_family="CT-rich"
repeat_region 22126..22382
                /rpt_family="CT-rich"
repeat_region complement(22420..22646)
                /rpt_family="AluY"
repeat_region complement(22653..22978)
                /rpt_family="MLTID"
repeat_region complement(22979..23285)
                /rpt_family="AluY"
repeat_region complement(23286..23475)
                /rpt_family="MLTID"
repeat_region 23488..23639
                /rpt_family="MER82"
repeat_region complement(23640..23944)
                /rpt_family="AluSx"
repeat_region complement(23948..24262)
                /rpt_family="AluJb"
repeat_region 24263..24550
                /rpt_family="MER82"
repeat_region 24616..24772
                /rpt_family="MER82"
repeat_region complement(25217..25305)
                /rpt_family="FAM"
repeat_region 25398..25478
                /rpt_family="LIME"
repeat_region 25555..25847
                /rpt_family="AluSx"
repeat_region complement(27234..27366)
                /rpt_family="FLAM_C"
repeat_region complement(27665..27965)
                /rpt_family="AluY"
repeat_region complement(28030..28320)
                /rpt_family="AluJb"
repeat_region complement(28328..28417)
                /rpt_family="MER113"
repeat_region complement(28458..28580)
                /rpt_family="AluSg/x"
repeat_region complement(28581..28888)
                /rpt_family="AluSg"
repeat_region complement(30425..30568)
                /rpt_family="L2"
repeat_region 31094..31139
                /rpt_family="CAGA)n"

```

```

repeat_region 31155..31294
                /rpt_family="GA-rich"
repeat_region 32987..33248
                /rpt_family="AluSg"
repeat_region 33249..33302
                /rpt_family="AluSg"
repeat_region 33398..33680
                /rpt_family="CA)n"
repeat_region 33710..34020
                /rpt_family="AluSx"
repeat_region 34027..34240
                /rpt_family="FLAM_C"
repeat_region 34243..34535
                /rpt_family="AluSg"
repeat_region 34536..34573
                /rpt_family="AT-rich"
repeat_region complement(35523..35699)
                /rpt_family="MER20"
repeat_region complement(35895..35992)
                /rpt_family="AluJb"
repeat_region 36356..36644
                /rpt_family="AluSg"
repeat_region 36845..37156
                /rpt_family="AluSg"
repeat_region complement(37214..37315)
                /rpt_family="Alu"
repeat_region complement(37408..37610)
                /rpt_family="MER20"
repeat_region 37625..37760
                /rpt_family="MIR"
repeat_region complement(38577..38873)
                /rpt_family="AluJb"
repeat_region complement(39109..39402)
                /rpt_family="AluSx"
repeat_region 39565..39768
                /rpt_family="LMB6"
repeat_region 39769..40073
                /rpt_family="AluSx"
repeat_region 40123..40228
                /rpt_family="GA-rich"
repeat_region 40238..40416
                /rpt_family="MER20"
exon           complement(41052..>41096)
                /gene="EEN"
                /number=1
repeat_region 41154..41250
                /rpt_family="CCG)n"
repeat_region 41305..41348
                /rpt_family="CCCCCG)n"

```

Query Match 7.2% Score 72; DB 9; Length 96308;
 Best Local Similarity 100.0%; Pred. No. 1e-30;
 Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 334 CGTCTCTACTAAATAACAAAATAGCCAGGTGGTGCCACACGCGCTGTAGTCCAGC 393
 |||||

Db 72056 CGTCTCTACTAAATAACAAAATAGCCAGGTGGTGCCACACGCGCTGTAGTCCAGC 71997
 |||||

Qy 394 TACTTGGGAGGC 405
 |||||

Db 71996 TACTTGGGAGGC 71985

RESULT 13

HS149A16

LOCUS

DEFINITION HSI49A16 173354 bp DNA linear PRI 12-DEC-1999
 Human DNA sequence from clone RPI-149A16 on chromosome 22 Contains
 an IGLC (Immunoglobulin Lambda Chain C) pseudogene, the RPL3 gene
 for Ret finger protein-like 3, the RPL3S gene for Ret finger
 protein-like 3 antisense, the gene for a novel Immunoglobulin
 Lambda Chain V family protein, the gene for a novel protein similar
 to mouse RGDs (RALGDS, RALGEF, Guanine Nucleotide Dissociation
 Stimulator A) and rabbit oncogene RSC, the gene for a novel protein


```

repeat_region 19818..19877
/note="MIR repeat: matches 69..128 of consensus"
repeat_region 20817..20983
/note="L2 repeat: matches 2276..2440 of consensus"
repeat_region 21858..21938
/note="L2 repeat: matches 2407..2488 of consensus"
repeat_region 21962..22065
/note="L2 repeat: matches 2595..2707 of consensus"
repeat_region 22311..22446
/note="68 copies 2 mer ac 61 conserved"
repeat_region 24135..24225
/note="MIR repeat: matches 26..118 of consensus"
misc_feature 24808..25247
/note="match: GSS: Em:AQ881697"
repeat_region 25356..25474
/note="L2 repeat: matches 2586..2710 of consensus"
repeat_region 25815..26103
/note="AluX repeat: matches 23..312 of consensus"
repeat_region 26197..26384
/note="MERG3B repeat: matches 1..436 of consensus"
repeat_region 26909..27216
/note="AluSg repeat: matches 1..286 of consensus"
repeat_region 28889..29892
/note="L2 repeat: matches 1496..2750 of consensus"
repeat_region 30213..30484
/note="L2 repeat: matches 2214..2502 of consensus"
repeat_region 30569..30651
/note="L2 repeat: matches 2665..2747 of consensus"
repeat_region 32066..32175
/note="MIRLJ repeat: matches 401..514 of consensus"
misc_feature complement(32187..32629)
/note="match: GSS: Em:AQ825726"
misc_feature complement(32796..33149)
/note="match: GSS: Em:AQ571968"
misc_feature 33171..33361
/note="match: STS: Em:F08813"
misc_feature 33242..33663
/note="match: GSS: Em:AQ153435"
repeat_region 34859..36399
/note="L1P2 repeat: matches 4602..6144 of consensus"
repeat_region 36506..36623
/note="MIR repeat: matches 59..187 of consensus"
repeat_region 37039..37108
/note="MER4D repeat: matches 485..557 of consensus"
repeat_region 37114..37291
/note="Harlequin repeat: matches 2301..2476 of consensus"

Query Match 7.2%; Score 72; DB 9; Length 173354;
Best Local Similarity 100.0%; Pred. No. 1.1e-30;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 232 CGGTGCTCACACCTGTATCCAGCAGCTTTGGAGGCTGAGAGGCTGATCAGCTGA 291
Db 110975 CGGTGCTCACACCTGTATCCAGCAGCTTTGGAGGCTGAGAGGCTGATCAGCTGA 111034

QY 292 AGTCAGGAGTTC 303
Db 111035 AGTCAGGAGTTC 111046

RESULT 14
AC011498/c
LOCUS
DEFINITION Homo sapiens chromosome 19 clone CTB-50L17, complete sequence.
AC011498
VERSION
KEYWORDS HTG.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.

```

```

TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 273403)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 273403)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (18-OCT-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Oct 18, 2001 this sequence version replaced gi:15022009.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 1.3.
FEATURES
source
1..273403
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTB-50L17"
BASE COUNT 59422 a 76105 c 73388 g 64488 t
ORIGIN
Query Match 7.2%; Score 72; DB 9; Length 273403;
Best Local Similarity 100.0%; Pred. No. 1.1e-30;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 CGTCTCTACTAAAATACAAAATAGCCAGGCTGTGGTGGCAGCGCTGTATGCCAGC 393
Db 62956 CGTCTCTACTAAAATACAAAATAGCCAGGCTGTGGTGGCAGCGCTGTATGCCAGC 62997

QY 394 TACTTGGGAGGC 405
Db 62896 TACTTGGGAGGC 62895

RESULT 15
AC024152/c
LOCUS
DEFINITION Homo sapiens chromosome 3 clone RP11-24E1, WORKING DRAFT SEQUENCE,
15 unordered pieces.
AC024152
VERSION AC024152.11 GI:11079270
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 162201)
AUTHORS Muzny,D.M., Adams,C., Ali-Othula,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Bivaga,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,X., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hoques,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,

```

Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,
 Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,
 Louis, H., Lozano, R., Lu, X., Lucier, A., Lucier, R., Luna, R.,
 Martinez, B., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M.,
 Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,
 Morgan, M., Morris, S., Moser, M., Neale, D., Newton, J., Newton, N.,
 Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokoko, S.,
 Ogih, M., Okwodu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,
 Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L.,
 Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojebokan, I., Rolfe, M.,
 Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoostari, N.,
 Siason, I., Sodergren, E., Sonaite, P., Sparks, A., Stanley, H.,
 Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K.,
 Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,
 Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R.,
 Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,
 Watlington, S., Williams, G., Williamson, A., Wleczek, R., Woodson, S.,
 Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 and Gibbs, R.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

Direct Submission
 Unpublished
 2 (bases 1 to 162201)
 Worley, K.C.

Submitted (25-FEB-2000) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 3, 2000 this sequence version replaced gi:94303818.

Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: HAHF

Center clone name: RP11-24E1

----- Summary Statistics

Sequencing vector: M13; L08821

Chemistry: Dye-terminator Big Dye; 95% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 137709 bases at least Q40

Consensus quality: 149473 bases at least Q40

Consensus quality: 154220 bases at least Q20

Estimated insert size: 15558; sum-of-contigs estimation

Quality coverage: 0x in Q20 bases; agarose-gel estimation

Quality coverage: 3.7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 15 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

1 39130: contig of 39130 bp in length

* 39131 56977: gap of unknown length

* 39231 56977: contig of 17467 bp in length

* 5698 56977: gap of unknown length

* 56798 74326: contig of 17529 bp in length

* 74327 74426: gap of unknown length

* 74427 90214: contig of 15788 bp in length

* 90215 90314: gap of unknown length

* 90315 106525: contig of 16211 bp in length

* 106526 106625: gap of unknown length

* 106626 119989: contig of 13364 bp in length

* 119990 120089: gap of unknown length

* 120090 130101: contig of 10012 bp in length

* 130102 130201: gap of unknown length

* 130202 138941: contig of 8740 bp in length

* 138942 139041: gap of unknown length

* 139042 145294: contig of 6253 bp in length

* 145295 145394: gap of unknown length

* 145395 150375: contig of 4881 bp in length

* 150376 150375: gap of unknown length

* 150377 150375: gap of unknown length

* 150378 155117: contig of 4642 bp in length

* 155118 157939: contig of 2822 bp in length

* 157940 158039: gap of unknown length

* 158040 159905: contig of 1866 bp in length

* 159906 160005: gap of unknown length

* 160006 161069: contig of 1064 bp in length

* 161070 161169: gap of unknown length

* 161170 162201: contig of 1032 bp in length.

FEATURES

source

1. 162201

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="3"

/clone="RP11-24E1"

BASE COUNT 44663 a 34049 c 34232 g 47778 t 1479 others

ORIGIN

Query Match 7.1% Score 71; DB 2; Length 162201;

Best Local Similarity 100.0%; Pred. No. 4.1e-30;

Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 GGATCACTTGAGCTGAGGCTTTGAGACCGCTGGCCACACGTCGAAACCCATCTC 204

Db 23363 GGATCACTTGAGCTGAGGCTTTGAGACCGCTGGCCACACGTCGAAACCCATCTC 204

QY 205 TACTAAAATA 215

Db 23303 TACTAAAATA 23293

RESULT 16

AC103560

LOCUS

DEFINITION

AC103560 AC022908

VERSION

AC103560.2 GI:21553246

KEYWORDS

HTG

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,

Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.

and Haugen, E.D.

TITLE

JOURNAL

REFERENCE

AUTHORS

Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,

Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.

and Haugen, E.D.

TITLE

JOURNAL

REFERENCE

AUTHORS

Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,

Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.

and Haugen, E.D.

TITLE

JOURNAL

COMMENT

On Jun 25, 2002 this sequence version replaced gi:17136120.

----- Genome Center

Center: University of Washington Genome Center

Center Code: UWGC

Web site: <http://www.genome.washington.edu>

Contact: uwgctg@u.washington.edu

Drafting Center: BCM

----- Project Information

Center project name: chr-3
Center clone name: RP11-673E20 (bc0526)

----- Summary Statistics

Sequencing vector: unknown; 53% of reads
Sequencing vector: plasmid; 47% of reads
Chemistry: Dye-terminator ET; 80% of reads
Chemistry: Dye-terminator Big Dye; 20% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 18599 bases at least Q40
Consensus quality: 186199 bases at least Q30
Consensus quality: 186239 bases at least Q20
Insert size: 186241; sum-of-contigs
Quality coverage: 9.7x in Q20 bases; sum-of-contigs

----- Overlapping Sequences:

5': Mapping in progress
3': RP11-252010 (UMGC:bc0324) AC024102

----- Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

----- Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

EcoRI				HindIII				BglII			
SeqDerMap	FrgrPrnt	SeqDerMap	FrgrPrnt	SeqDerMap	FrgrPrnt	SeqDerMap	FrgrPrnt	SeqDerMap	FrgrPrnt	SeqDerMap	FrgrPrnt
8696	8757	2175	2179	3135	3386						
6	<800	6382	6493	2067	1978						
1052	1002	512	<800	5815	5770						
17118	18865	449	<800	6118	6083						
9222	8960	2668	2671	1999	1978						
45	<800	1406	1394	5551	5770						
862	886	1120	1137	3666	3646						
982	1002	166	<800	221	<800						
2317	2306	6933	6956	2789	2866						
7842	7838	4355	4624	1175	1171						

1417	1383	4893	4874	2090	1978
3428	3541	8783	8743	4135	4303
9648	9521	3313	3337	2904	2866
1701	1691	657	<800	1627	1594
157	<800	1892	1892	751	758
2197	2306	4295	4303	8173	8038
3444	3362	1900	1892	948	958
2946	3189	3426	3647	3399	3646
1267	1255	3682	3925	1976	1978
999	1002	10582	10486	7002	6945
558	<800	3894	4184	10589	10321
3330	3189	649	<800	1130	1139
2257	2306	565	<800	1638	1594
1095	1067	298	<800	750	758
887	886	3312	3337	7554	7330
2276	2306	566	<800	382	<800
2533	2665	544	<800	45	<800
2593	2665	8816	8743	2967	3097
993	1002	1717	1711	5871	5770
1403	1383	1620	1547	3777	4021
1912	1876	343	<800	868	958
9030	8960	1144	1137	2198	2178
339	<800	457	<800	6062	5770
6659	6780	9478	9430	937	958
6650	6780	9092	9157	1764	1749
7426	7503	2445	2462	950	958
1570	1514	148	<800	984	958
1506	1514	4668	4874	520	<800
2275	2306	10544	10486	1650	1594
10930	10522	6248	6209	1338	1341
995	1002	7286	7252	1957	1978
737	<800	4827	4874	3561	3646
6846	6939	961	957	3658	3646
4499	4491	12517	12544	2545	2581
268	<800	6185	6209	4378	4611
3112	3189	572	<800	5041	5391

Query Match 7.1%; Score 71; DB 9; Length 186241;
 Best Local Similarity 100.0%; Pred. No. 4.2e-30;
 Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 GGATCATTGAGTGGAGTTTGGACCCAGCCTGGCCACACGCTGAAACCCCATCTC 204
 |||||
 DB 28955 GGATCATTGAGTGGAGTTTGGACCCAGCCTGGCCACACGCTGAAACCCCATCTC 29014
 |||||

QY 205 TACTAAATA 215
 |||||
 DB 29015 TACTAAATA 29025
 |||||

RESULT 17
 AL161445/c
 LOCUS
 DEFINITION Human DNA sequence from clone Rpl1-326F20 on chromosome 9 Contains
 the GGTAL gene for glycoprotein alpha-galactosyl transferase 1, the
 SPINK4 gene for Kazal type 4 serine protease inhibitor, the 3', and
 of the BAG1 gene for BCL2-associated athanogene and a CpG island,
 complete sequence.

ACCESSION AL161445
 VERSION AL161445.10 GI:11137678
 KEYWORDS HTG; athanogene; BAG1; BCL2; CpG island; galactosyltransferase;
 GGTAL; glycoprotein; Kazal; serine protease inhibitor; SPINK4.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 163396)

REFERENCE
 AUTHORS Ransay, H.
 TITLE Direct Submission
 JOURNAL Submitted (04-DEC-2001) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humcrquest@sanger.ac.uk
 humcrquest@sanger.ac.uk
 On Nov 10, 2000 this sequence version replaced gi:11071610.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPEP; Information
 on the WORMPEP database can be found at
 http://www.sanger.ac.uk/Projects/C.elegans/wormpep
 This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
 Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr9
 This sequence is the entire insert of clone Rpl1-326F20 The true

left end of clone Rpl1-344B24 is at 100362 in this sequence. The
 true right end of clone Rpl1-54K16 is at 54767 in this sequence.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. Rpl1-326F20 is from
 the library RPI-11.2 constructed by the group of Pieter de Jong.
 For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pBAC3.6.

FEATURES
 Location/Qualifiers
 1..163396
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="9"
 /clone_lib="RPI-326F20"
 /clone_lib="RPI-11.2"
 5..237
 /note="AluY repeat: matches 77. .309 of consensus"
 239..314
 /note="L1M4 repeat: matches 2591. .2666 of consensus"
 331..438
 /note="MER44A repeat: matches 225. .331 of consensus"
 440..555
 /note="AluSq/x repeat: matches 11. .126 of consensus"
 566..708
 /note="AluJo/FRAM repeat: matches 158. .298 of consensus"
 795..1193
 /note="L1M4 repeat: matches 3657. .4070 of consensus"
 1194..1491
 /note="AluSp repeat: matches 1. .296 of consensus"
 1492..1680
 /note="L1M4 repeat: matches 4070. .4265 of consensus"
 1681..1981
 /note="AluX repeat: matches 1. .301 of consensus"
 1982..2181
 /note="L1M4 repeat: matches 4265. .4471 of consensus"
 2182..2486
 /note="AluX repeat: matches 1. .304 of consensus"
 2487..2783
 /note="L1M4 repeat: matches 4471. .4701 of consensus"
 3046..3192
 /note="AluSq/x repeat: matches 1. .147 of consensus"
 3196..3247
 /note="MIR repeat: matches 17. .66 of consensus"
 4038..4566
 /note="MER41B repeat: matches 150. .635 of consensus"
 4567..4862
 /note="AluSq repeat: matches 1. .300 of consensus"
 4863..5008
 /note="MER41B repeat: matches 1. .150 of consensus"
 6433..6490
 /note="MIR repeat: matches 81. .138 of consensus"
 7523..7596
 /note="L1P1 repeat: matches 5753. .5825 of consensus"
 8418..8461
 /note="L1MB4 repeat: matches 6132. .6184 of consensus"
 8482..8769
 /note="AluX repeat: matches 1. .304 of consensus"
 8770..8783
 /note="L1MB4 repeat: matches 6145. .6132 of consensus"
 8784..9085
 /note="AluSp repeat: matches 1. .302 of consensus"
 9086..9106
 /note="L1MB4 repeat: matches 6122. .6146 of consensus"
 9152..9751
 /note="L1MB3 repeat: matches 5512. .6125 of consensus"
 10464..10534
 /note="MIR repeat: matches 61. .130 of consensus"
 11734..11790

```
/note="L2 repeat: matches 2688. .2742 of consensus"
11747. .11869
/note="MIR repeat: matches 132. .252 of consensus"
11874. .12110
/note="FLAM_C repeat: matches 1. .134 of consensus"
12121. .12219
/note="MIR repeat: matches 35. .142 of consensus"
12243. .12346
/note="52 copies 2 mer tg 71% conserved"
12254. .12343
/note="9 copies 10 mer gtgtgtgtgt 72% conserved"
complement(12461. .12790)
/note="match: GSS: Em:AQ210501"
complement(12571. .12678)
/note="match: GSS: Em:AQ177335"
12602. .13170
/note="MER21B repeat: matches 232. .794 of consensus"
13171. .13289
/note="FLAM_A repeat: matches 1. .122 of consensus"
13290. .13517
/note="MER21B repeat: matches 4. .232 of consensus"
14033. .14159
/note="MIR repeat: matches 112. .257 of consensus"
complement(14283. .70891)
/gene="GGTA1"
complement(join(14283. .17232,17420. .17524,19637. .19759,
24065. .24252,38835. .39070,70404. .70891))
/gene="GGTA1"
/product="ba326F20.1 (glycoprotein, alpha-galactosyl
transferase 1)"
/note="glycoprotein, alpha-galactosyl transferase 1
match: CDNAS: Em:J03980 Em:D00314 Em:E02227 Em:M13569
Em:X14085 Em:A23697 Em:X14558 Em:M22921 Em:U10473
Em:R02228 Em:X13223 Em:D29805 Em:M13214 Em:X55415
Em:U19890 Em:Y12510 Em:M13701 Em:U19889 Em:AF038660
Em:U10474 Em:AB024434 Em:AB019541 Em:U10472 Em:AF142670
match: ESTs: Em:BF124086 Em:AW467118 Em:AA769459
Em:AA476930 Em:AI192425 Em:AI572046 Em:AI149377
Em:AA830559 Em:AA889902 Em:AI028425 Em:AW362033
Em:AI149203 Em:AI565973"
/evidence=not_experimental
complement(14283)
/gene="GGTA1"
complement(14305. .14310)
/gene="GGTA1"
14897. .14934
/note="19 copies 2 mer aa 92% conserved"
complement(15755. .17099)
/gene="GGTA1"
/note="match: STS: Em:G05861"
complement(16658. .16912)
/gene="GGTA1"
/note="match: STS: Em:G59840"
complement(join(17100. .17232,17420. .17524,19637. .19759,
24065. .24252,38835. .39070,70404. .70815))
/gene="GGTA1"
/note="match: proteins: Tr:O60512"
/codon_start=1
/evidence=not_experimental
/product="ba326F20.1 (glycoprotein, alpha-galactosyl
transferase 1)"
/protein_id="CAD13306.1"
/db_xref="GI:17384434"
/translation="MRLREPLLSGAAPGASIQRACLRLLVAVCALHLGVILVYLAG
RLSLRLPOLGVSTPLQGSNSAAIGQSGELRTGGARPPPLGASPPRGDSSP
VVDSCPGASNLTVPEVHTTALSIPACPEESPLLPGMLTEPMVDLELVAKQNP
VWNGRTAPRDCVSPKVALIIPNRRQELKILWLYHPVQLQKQDLCIYVINAQF
DTFNRAKLLNVGFQALKDYDTCFVFSVDLIPMNDHNAYRCFSQPHRISVAMDKF
GSLPYVGFQGSALSQQQLTINGFPNNWGGGDEDDIFNRLVFRGMSISRNAY
VGRCRMIRHSRDKKNEPNPQRFRIATKETMLSDGLNSLYQVLDVQRYPLTYITV
DIGTFS"
17946. .18109
/note="L2 repeat: matches 2241. .2411 of consensus"
```

```
repeat_region 19126. .19243
/note="L1MC5 repeat: matches 7805. .7927 of consensus"
19877. .20157
/note="AluSc repeat: matches 11. .299 of consensus"
20168. .20468
/note="AluSx repeat: matches 1. .301 of consensus"
22027. .22341
/note="AluJo repeat: matches 4. .309 of consensus"
22507. .22807
/note="AluSx repeat: matches 1. .301 of consensus"
23081. .23376
/note="AluJo repeat: matches 1. .297 of consensus"
23549. .23852
/note="AluJb repeat: matches 1. .305 of consensus"
24409. .24625
/note="L1ME repeat: matches 5615. .5826 of consensus"
25003. .25292

Query Match 7.0%; Score 70; DB 9; Length 163396;
Best Local Similarity 100.0%; Pred. No. 1.6e-29;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 340 TACTAAATAACAAAAATTAGCCAGGTGTGGCCACACGCTGTAGTCCAGCTACTTG 399
|||||
Db 20352 TACTAAATAACAAAAATTAGCCAGGTGTGGCCACACGCTGTAGTCCAGCTACTTG 20293
|||||

QY 400 GGAGGCTGAG 409
|||||
Db 20292 GGAGGCTGAG 20283

RESULT 18
AC069173/c 197143 bp DNA linear HTG 04-OCT-2000
LOCUS Homo sapiens chromosome 3 clone RP11-594G13 map 3, WORKING DRAFT
DEFINITION AC069173.2 GI:10567947
ACCESSION AC069173
VERSION AC069173.2
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 197143)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bida,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,S.,
Campopiano,A., Castelle,A., Choquel,Y., Colangelo,M., Collins,S.,
Collins,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S.,
Dodgson,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Gallagan,J., Gardyna,S., Glinde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehotzky,J.,
Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisan,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
TITLE
JOURNAL
Submitted (21-MAY-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Oct 4, 2000 this sequence version replaced gi:7960339.
```

All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center -----
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIGR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information -----
Center project name: L10393
Center clone name: 594.G.L13
----- Summary Statistics -----
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 181889 bases at least Q40
Consensus quality: 190399 bases at least Q30
Consensus quality: 192365 bases at least Q20
Insert size: 205000; agarose-fp
Insert size: 194343; sum-of-contigs
Quality coverage: 3.8 in Q20 bases; agarose-fp
Quality coverage: 4.0 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 29 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	4626:	contig of 4626 bp in length	
*	4627	4726:	gap of 100 bp
*	4727	6065:	contig of 1339 bp in length
*	6066	6165:	gap of 100 bp
*	6166	7387:	contig of 1222 bp in length
*	7388	7487:	gap of 100 bp
*	7488	9199:	contig of 1712 bp in length
*	9200	9299:	gap of 100 bp
*	9300	12161:	contig of 2862 bp in length
*	12162	12261:	gap of 100 bp
*	12262	14008:	contig of 1747 bp in length
*	14009	14108:	gap of 100 bp
*	14109	17158:	contig of 3050 bp in length
*	17159	17258:	gap of 100 bp
*	17259	21108:	contig of 3850 bp in length
*	21109	21208:	gap of 100 bp
*	21209	25235:	contig of 4047 bp in length
*	25256	25355:	gap of 100 bp
*	25356	29225:	contig of 3870 bp in length
*	29226	29325:	gap of 100 bp
*	29326	34349:	contig of 5024 bp in length
*	34350	34449:	gap of 100 bp
*	34450	38902:	contig of 4453 bp in length
*	38903	39002:	gap of 100 bp
*	39003	45778:	contig of 6776 bp in length
*	45779	45878:	gap of 100 bp
*	45879	51119:	contig of 5241 bp in length
*	51120	51219:	gap of 100 bp
*	51220	55931:	contig of 4712 bp in length
*	55932	56031:	gap of 100 bp
*	56032	60610:	contig of 4579 bp in length
*	60611	60710:	gap of 100 bp
*	60711	67344:	contig of 6634 bp in length
*	67345	67444:	gap of 100 bp
*	67445	71789:	contig of 4345 bp in length
*	71790	71889:	gap of 100 bp
*	71890	78631:	contig of 6742 bp in length
*	78632	78731:	gap of 100 bp
*	78732	86021:	contig of 7290 bp in length
*	86022	86121:	gap of 100 bp
*	86122	92375:	contig of 6254 bp in length
*	92376	92475:	gap of 100 bp

```

misc_feature 121415..134427
              /note="assembly_fragment"
misc_feature 134528..147968
              /note="assembly_fragment"
misc_feature 148069..161342
              /note="assembly_fragment"
misc_feature 161443..177031
              /note="assembly_fragment"
misc_feature 177132..197143
              /note="assembly_fragment"
BASE COUNT 49796 a 47285 c 47847 g 49412 t 2803 others
ORIGIN
Query Match      7.0%; Score 70; DB 2; Length 197143;
Best Local Similarity 100.0%; Pred. No. 1.6e-29;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 CGTCTCTACTAAATACAAAATTAGCCAGGTGTGGTGCACACGCTGTAGTCCGAGC 393
      |||
Db 155629 CGTCTCTACTAAATACAAAATTAGCCAGGTGTGGTGCACACGCTGTAGTCCGAGC 155570

QY 394 TACTTGGGAG 403
      |||
Db 155569 TACTTGGGAG 155560

RESULT 19
AC023603/c
LOCUS
DEFINITION Homo sapiens chromosome 3p clone RP11-330614, WORKING DRAFT
ACCESSION AC023603
VERSION AC023603.1 GI:6980169
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 170513)
AUTHORS Zhao,Y., Zhang,C., Lin,W., Dong,H., Wan,M., Xu,S., Gu,W., Tu,Y.,
          Jia,J., Wu,C., Lu,G., Zhong,M., Zhou,Y., Ren,S., Fu,G., Chen,Z. and
          Huang,M.
          Direct Submission
          Submitted (16-FEB-2000) Genomic Dept., Chinese National Human
          Genome Center at Shanghai, 315 Guo Shou ying Road, Shanghai,
          Shanghai 201203, P. R. China
          -----Genome Center Informations-----
          Center: Chinese National Human Genome Center at Shanghai
          Center Code: CHGC
          Web site: http://www.chgc.sh.cn
          Email: mhuan@chgc.sh.cn or fugang@chgc.sh.cn
          -----End Genome Center Informations-----
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 14 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *
          * 1 2230: contig of 2230 bp in length
          * gap of unknown length
          * 2231 4248: contig of 2018 bp in length
          * gap of unknown length
          * 4249 9512: contig of 5264 bp in length
          * gap of unknown length
          * 9513 14048: contig of 4536 bp in length
          * gap of unknown length
          * 14049 18094: contig of 4045 bp in length
          * gap of unknown length
          * 18095 23536: contig of 5442 bp in length
          * gap of unknown length
          * 23537 29239: contig of 5703 bp in length

```

```

* 29240 39461: gap of unknown length
* contig of 10222 bp in length
* 39462 49383: contig of 9922 bp in length
* gap of unknown length
* 49384 62126: contig of 12743 bp in length
* gap of unknown length
* 62127 84920: contig of 22794 bp in length
* gap of unknown length
* 84921 108776: contig of 23856 bp in length
* gap of unknown length
* 108777 135846: contig of 27070 bp in length
* gap of unknown length
* 135847 170513: contig of 34667 bp in length.
FEATURES
source
1. 170513
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="3p"
   /clone="RP11-330614"
BASE COUNT 53252 a 32291 c 31832 g 53096 t 42 others
ORIGIN
Query Match      6.7%; Score 67; DB 2; Length 170513;
Best Local Similarity 100.0%; Pred. No. 9.8e-28;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 340 TACTAAATACAAAATTAGCCAGGTGTGGTGCACACGCTGTAGTCCGACTACTTG 399
      |||
Db 119984 TACTAAATACAAAATTAGCCAGGTGTGGTGCACACGCTGTAGTCCGACTACTTG 119925

QY 400 GGAGGCT 406
      |||
Db 119924 GGAGGCT 119918

RESULT 20
AC006042/c
LOCUS
DEFINITION Homo sapiens BAC clone RP11-505D17 from 7p22-p21, complete
          sequence.
ACCESSION AC006042
VERSION AC006042.2 GI:4508120
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 178361)
AUTHORS Sulston,J.E. and Waterston,R.
          TITLE Toward a complete human genome sequence
          JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
          MEDLINE 99063792
          PUBMED 9847074
REFERENCE 2 (bases 1 to 178361)
AUTHORS Hou,S., Wohlman,P. and Le,T.
          TITLE The sequence of Homo sapiens BAC clone RP11-505D17
          JOURNAL Unpublished
          REFERENCE 3 (bases 1 to 178361)
          AUTHORS Waterston,R.H.
          TITLE Direct Submission
          JOURNAL Submitted (22-NOV-1998) Genome Sequencing Center, Washington
          University School of Medicine, 4444 Forest Park Parkway, St. Louis,
          MO 63108, USA
          REFERENCE 4 (bases 1 to 178361)
          AUTHORS Waterston,R.H.
          TITLE Direct Submission
          JOURNAL Submitted (24-MAR-1999) Genome Sequencing Center, Washington
          University School of Medicine, 4444 Forest Park Parkway, St. Louis,
          MO 63108, USA
          REFERENCE 5 (bases 1 to 178361)
          AUTHORS Waterston,R.
          TITLE Direct Submission

```

JOURNAL

Submitted (19-AUG-1999) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
6 (bases 1 to 178361)
Waterston,R.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Direct Submission
Submitted (21-DEC-1999) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Mar 24, 1999 this sequence version replaced gi:3980487.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc>

Contact: sapiens@watson.wustl.edu

----- Summary Statistics

----- Center project name: H_NH0505D17

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhri.nih.gov/DIR/GB/CHR7/>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frenken, E., Tatenoe, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)
VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-438H20; the clone sequenced to the right is RP11-560C1. The actual start of this clone is at base position 1 of RP11-505D17; actual end is at base position 178361 of RP11-505D17.

FEATURES

Source

Location/Qualifiers

1..178361
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7p22-p21"
/clone="RP11-505D17"
/clone_lib="RPCI-11"
18..267

repeat_region

/rpt_family="L2"

repeat_region

/rpt_family="L2"

repeat_region

/rpt_family="(CA)n"

repeat_region

817..1043
/rpt_family="L1"

misc_feature

838..1267

/note="match to EST AI025912 (NID:g3241525) ov95e10.sl"

repeat_region

/rpt_family="Alu"

misc_feature

2041..2511

repeat_region

/note="match to EST AA759276 (NID:g2807139) ah90e10.sl"

repeat_region

5336..5468
/rpt_family="MIR"

repeat_region

6287..6449
/rpt_family="L1"

repeat_region

7747..8030
/rpt_family="Alu"

repeat_region

8032..8058
/rpt_family="(TAAAA)n"

misc_feature

8237..8585
/note="match to EST AI223166 (NID:g3805369) qg53d05.xl"

misc_feature

8241..8581
/note="match to EST AA770501 (NID:g2821739) ah74b12.sl"

repeat_region

8248..8649
/note="match to EST AA926996 (NID:g3075993) om26d02.sl"

repeat_region

9367..9603
/rpt_family="MaLR"

repeat_region

9604..9896
/rpt_family="Alu"

repeat_region

9897..10077
/rpt_family="MaLR"

repeat_region

10680..10981
/rpt_family="Alu"

repeat_region

10998..11037
/rpt_family="(CA)n"

repeat_region

11186..11220
/rpt_family="AT-rich"

repeat_region

11587..11713
/rpt_family="MIR"

repeat_region

11834..12132
/rpt_family="Alu"

repeat_region

13508..13607
/rpt_family="(GAAA)n"

repeat_region

13611..13902
/rpt_family="Alu"

repeat_region

14757..14932
/rpt_family="MIR"

repeat_region

16538..16686
/rpt_family="MERL_type"

repeat_region

16819..17280
/rpt_family="MaLR"

repeat_region

17727..18009
/rpt_family="Alu"

repeat_region

18104..18313
/rpt_family="L2"

repeat_region

18487..18974
/rpt_family="L2"

repeat_region

18975..19361
/rpt_family="Alu"

repeat_region

19262..20642
/rpt_family="L2"

repeat_region

21084..21133
/rpt_family="A-rich"

repeat_region

21186..21766
/rpt_family="L2"

repeat_region

23013..23147
/rpt_family="Alu"

repeat_region

23416..23768
/rpt_family="L2"

repeat_region

25129..25424
/rpt_family="Alu"

repeat_region

25553..25603
/rpt_family="AT-rich"

repeat_region

25755..25989
/rpt_family="MIR"

misc_feature

26095..26418
/note="match to EST AI302527 (NID:g3961873) qn47g06.xl"

misc_feature

26132..26418
/note="match to EST AI271704 (NID:g3890871) qj87f01.xl"

misc_feature

26865..28262

/note="CpG_island (%GC=73.1, o/e=0.80, *CpGs=164)"

repeat_region 27225..27245
 /rpt_family="CGG)n"
 repeat_region 27656..27758
 /rpt_family="L1"
 repeat_region 27759..27778
 /rpt_family="(GGA)n"
 repeat_region 27779..28238
 /rpt_family="L1"
 misc_feature 28108..28205
 /note="match to EST N32168 (NID:g1152567) YV23f11.s1"
 gene <28109..144935
 /gene="WUGSC:H.NH0505D17.1"
 CDS join(<28109..28205,62305..62456,80880..80966,
 113833..113946,118493..118645,129318..129528,
 143294..143414,144590..144935)
 /gene="WUGSC:H.NH0505D17.1"
 /note="supported by human ESTs
 AI681256.1(NID:g4891438),N32168.1(NID:g1152567), and

Query Match 6.7%; Score 67; DB 9; Length 178361;
 Best Local Similarity 100.0%; Pred. No. 9.8e-28;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 340 TACTAAAAATACAAAATAGCAGGTGGTGGCACACGCTGTAGTCCAGCTACTTG 399
 |||||
 Db 39904 TACTAAAAATACAAAATAGCAGGTGGTGGCACACGCTGTAGTCCAGCTACTTG 39845

QY 400 GGAGGCT 406
 |||||
 Db 39844 GGAGGCT 39838

RESULT 21
 AL731576/c
 LOCUS Human DNA sequence from clone RP11-178C16 on chromosome 10,
 complete sequence.
 DEFINITION AL731576 AC025425
 ACCESSION AL731576.9 GI:22002710
 VERSION HTG.
 KEYWORDS Homo sapiens.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 AUTHORS Taylor,S.
 TITLE Direct Submission
 JOURNAL Submitted (26-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquery@sanger.ac.uk
 COMMENT On Jul 29, 2002 this sequence version replaced gi:21322404.
 Draft Sequence Produced by Genome Therapeutics Corp, 100 Beaver
 Street, Waltham, MA 02453, USA
 http://www.genomecorp.com

During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30);
 an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Enr., EMBL; Sw.,
 SWISSPROT; Tr., TrEMBL; Wp., WormPEP; Information on the WormPEP
 database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence

was generated from part of bacterial clone contigs of human
 chromosome 10, constructed by the Sanger Centre Chromosome 10
 Mapping Group. Further information can be found at
 http://www.sanger.ac.uk/AGP/Chr10

RP11-178G16 is from the library RPCR-11.1 constructed by the group
 of Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pBACE3.6

----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquery@sanger.ac.uk

FEATURES

source 1..104656
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="10"
 /clone="RP11-178G16"
 /clone_lib="RPCI-11.1"
 BASE COUNT 29430 a 21161 c 22634 g 31431 t
 ORIGIN

Query Match. 6.6%; Score 66; DB 9; Length 104656;
 Best Local Similarity 100.0%; Pred. No. 3.7e-27;
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 335 GTCCTACTAAAAATACAAAATAGCAGGTGGTGGCACACGCTGTAGTCCAGCT 394
 |||||
 Db 72280 GTCCTACTAAAAATACAAAATAGCAGGTGGTGGCACACGCTGTAGTCCAGCT 72221

QY 395 ACTTGG 400
 |||||

Db 72220 ACTTGG 72215

RESULT 22
 AY052369/c

LOCUS Homo sapiens PP2A B56 gamma gene, complete cds, alternative splice
 products.
 DEFINITION AY052369
 ACCESSION AY052369.1 GI:16303629
 VERSION
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 AUTHORS Muneer,S., Ramalingam,V., Wyatt,R., Schultz,R.A., Minna,J.D. and
 Kamibayashi,C.

TITLE Genomic organization and mapping of the gene encoding the PP2A
 B56gamma regulatory subunit
 JOURNAL Genomics 79 (3), 344-348 (2002)
 MEDLINE 21853305
 PUBMED 11863364

REFERENCE 2 (bases 1 to 121028)
 Muneer,S., Kamibayashi,C. and Minna,J.D.
 DIRECT SUBMISSION
 TITLE Direct Submission
 JOURNAL Submitted (24-AUG-2001) Hamon Center for Therapeutic Oncology,
 University of Texas Southwestern Medical Center, 6000 Harry Hines
 Blvd, Dallas, TX 75390, USA

FEATURES
 source 1..121028
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="14q32.2"

mRNA
 join(2507..2643,49292..49491,74767..74877,75863..75955,
 76038..76168,82840..82899,85603..85711,87110..87163,
 94322..94492,99005..99132,102192..102293,105004..105076,
 110434..110550,117745..120289)

```

mRNA
/product="pp2A B56 gamma 3"
join(2507..2643,49292..49491,74767..74877,75863..75955,
76038..76168,82840..82899,85603..85711,87110..87163,
94322..94492,99005..99132,102192..102293,105004..105076,
117745..1120289)
/product="pp2A B56 gamma 2"
join(2507..2643,49292..49491,74767..74877,75863..75955,
76038..76168,82840..82899,85603..85711,87110..87163,
94322..94492,99005..99132,102192..102293,105004..105226)
/product="pp2A B56 gamma 1"
2507..2643
/number=1
join(2580..2643,49292..49491,74767..74877,75863..75955,
76038..76168,82840..82899,85603..85711,87110..87163,
94322..94492,99005..99132,102192..102293,105004..105076,
110434..110550,117745..117876)
/note="protein phosphatase; splice variant 3"
/codon_start=1
/product="pp2A B56 gamma 3"
/protein_id="AAL14779.1"
/db_xref="GI:16303632"
/translation="MVVDAANSNGFPQPVLLHIRDVPPADQEKLFQIKLRQCCLFD
FVSDPLDLKWEKRAALSEMVEYITHNRNVTETIPEVNHFAVNMFTLPPSSN
PTGAEDPEDEPTLEAAWPHLOLVYEFELFLESDFDPQNTAKYIDQKVLQLEL
FSDSDPRDFLKTTHRIYKGLRAYIRKQINNIYFRIYETEHNGIAELLEIL
GSILINGFALPLKEHKIFLLKVLPLHKVLSVYHPQALCVVQFLEKDSITLTPVY
MALKYWPKTHSPKEVNFLELEIDVIEPSEFVKIMEPLFROLAKCVSSPHFOVAE
RALIYWNNEYIMSLISDNAKILPIMFPSLYRNKSTHNKTIHGLIYNALKLFEMNQ
KLFDCTQOQFAEKLKEKLKMKEREAWKIEENLAKANPQTVVYSAQSMSTIPVAMET
DPLFEDVQMLRKTVDKHAQKQPKKDRPLAKRSELPODPHTKKALEAHCRADEL
ASQDGR"
join(2580..2643,49292..49491,74767..74877,75863..75955,
76038..76168,82840..82899,85603..85711,87110..87163,
94322..94492,99005..99132,102192..102293,105004..105076,
117745..117876)
/note="protein phosphatase; splice variant 2"
/codon_start=1
/product="pp2A B56 gamma 2"
/protein_id="AAL14778.1"
/db_xref="GI:16303631"
/translation="MVVDAANSNGFPQPVLLHIRDVPPADQEKLFQIKLRQCCLFD
FVSDPLDLKWEKRAALSEMVEYITHNRNVTETIPEVNHFAVNMFTLPPSSN
PTGAEDPEDEPTLEAAWPHLOLVYEFELFLESDFDPQNTAKYIDQKVLQLEL
FSDSDPRDFLKTTHRIYKGLRAYIRKQINNIYFRIYETEHNGIAELLEIL
GSILINGFALPLKEHKIFLLKVLPLHKVLSVYHPQALCVVQFLEKDSITLTPVY
MALKYWPKTHSPKEVNFLELEIDVIEPSEFVKIMEPLFROLAKCVSSPHFOVAE
RALIYWNNEYIMSLISDNAKILPIMFPSLYRNKSTHNKTIHGLIYNALKLFEMNQ
KLFDCTQOQFAEKLKEKLKMKEREAWKIEENLAKANPQADPKDRPLAKRSEL
PODPHTKKALEAHCRADELASQDGR"
join(2580..2643,49292..49491,74767..74877,75863..75955,
76038..76168,82840..82899,85603..85711,87110..87163,
94322..94492,99005..99132,102192..102293,105004..105100)
/note="protein phosphatase; splice variant 1"
/codon_start=1
/product="pp2A B56 gamma 1"
/protein_id="AAL14777.1"
/db_xref="GI:16303630"
/translation="MVVDAANSNGFPQPVLLHIRDVPPADQEKLFQIKLRQCCLFD
FVSDPLDLKWEKRAALSEMVEYITHNRNVTETIPEVNHFAVNMFTLPPSSN
PTGAEDPEDEPTLEAAWPHLOLVYEFELFLESDFDPQNTAKYIDQKVLQLEL
FSDSDPRDFLKTTHRIYKGLRAYIRKQINNIYFRIYETEHNGIAELLEIL
GSILINGFALPLKEHKIFLLKVLPLHKVLSVYHPQALCVVQFLEKDSITLTPVY
MALKYWPKTHSPKEVNFLELEIDVIEPSEFVKIMEPLFROLAKCVSSPHFOVAE
RALIYWNNEYIMSLISDNAKILPIMFPSLYRNKSTHNKTIHGLIYNALKLFEMNQ
KLFDCTQOQFAEKLKEKLKMKEREAWKIEENLAKANPQVLKRLIT"
/number=2
74767..74877
/number=3
75863..75955
/number=4
76038..76168
/number=5

```

```

exon      82840..82899
/number=6
exon      85603..85711
/number=7
exon      87110..87163
/number=8
exon      94322..94492
/number=9
exon      99005..99132
/number=10
exon      102192..102293
/number=11
exon      105004..105076
/number=12
exon      105077..105226
/number=12A
exon      110434..110550
/number=13
exon      117745..1120289
/number=14
BASE COUNT 31520 a 26430 c 27189 g 35887 t 2 others
ORIGIN
Query Match      6.6%   Score 66;   DB 9;   Length 121028;
Best Local Similarity 100.0%;   Pred. No. 3.8e-27;
Matches 66;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;
QY 142 GGTGGATCACTTGAGTGGTCTGAGTCTTTGAGACCAAGCTGCGCAACACGGTGAACCCCAT 201
      |||||||
DB 26660 GGTGGATCACTTGAGTGGTCTGAGTCTTTGAGACCAAGCTGCGCAACACGGTGAACCCCAT 26601
      |||||||
QY 202 CTCCTAC 207
DB 26600 CTCCTAC 26595
RESULT 23
AC110023/c
LOCUS      AC110023      166975 bp      DNA      linear      PRI 11-APR-2002
DEFINITION Homo sapiens chromosome 15, clone RP11-739G5, complete sequence.
ACCESSION      AC110023
VERSION      AC110023.3      GI:20043183
KEYWORDS      HTG.
SOURCE      Homo sapiens.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
1 (bases 1 to 166975)
Homo sapiens chromosome 15, clone RP11-739G5
Unpublished
2 (bases 1 to 166975)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,I., Bouckgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McKean,P., McKernan,K., Meidrim,J., Meneus,L., Mihova,T.,
Mlenga,C., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuppback,R., Seaman,S.,
Severy,P., Spencer,N., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,

```



```

repeat_region      /rpt_family="MERS8A"
16310..16353
repeat_region      /rpt_family="MIR"
17063..17083
repeat_region      complement(17197..17482)
repeat_region      /rpt_family="AluSc"
complement(17483..17626)
repeat_region      /rpt_family="MIR"
complement(18296..18429)
repeat_region      /rpt_family="MIR"
complement(20132..20502)
repeat_region      /rpt_family="MIRII"
21039..21062
repeat_region      /rpt_family="AT_rich"
21276..21302
repeat_region      /rpt_family="AT_rich"
21896..22065
repeat_region      /rpt_family="L1PB3"

Query Match          6.6%; Score 66; DB 9; Length 166975;
Best Local Similarity 100.0%; Pred. NO. 3.8e-27;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 346 AAATACAAAATTAGCCAGGTGTGTGGCACAGCGCTAGTCCAGCTACTTGGGAGGC 405
|||||
Db 23765 AAATACAAAATTAGCCAGGTGTGTGGCACAGCGCTAGTCCAGCTACTTGGGAGGC 23706

QY 406 TGAGGC 411
|||||
Db 23705 TGAGGC 23700

RESULT 24
AC024262          182627 bp      DNA      linear      HTG 28-MAR-2001
LOCUS            Homo sapiens chromosome 15 clone RP11-164C12, WORKING DRAFT
SEQUENCE, 4 unordered pieces.
AC024262
AC024262.17 GI:13470160
HTG; HTGS-PHASE1; HTGS-DRAFT; HTGS-FULLTOP.
Homo sapiens.
ORGANISM         Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 182627)
Fedorov, N., Glukhov, S., Hansen, N., Herman, Z.S., Hyman, R.,
Mao, J., Komp, C., Kottler, S., Lam, B., Marathe, R., Miranda, M.,
Morehouse, A.J., Nguyen, M., Oefner, P., Palm, C.J., Ramirez, D.,
Southwick, A.M., Webb, C., Wilhelmy, J., Yu, S. and Davis, R.W.
Unpublished
2 (bases 1 to 182627)
Bruno, D., Conn, L., Della Rosa, M., Faulkner, D., Federspiel, N.,
Glukhov, S., Hansen, N., Hyman, R., Mao, J., Marathe, R.,
Morehouse, A.J., Oefner, P., Palm, C.J., Ramirez, D., Wilhelmy, J.,
Yu, S. and Davis, R.W.
Direct Submission
Submitted (28-FEB-2000) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
COMMENT          On Mar 28, 2001 this sequence version replaced gi:13450006.
----- Genome Center
Center: Stanford DNA Sequencing and Technology Development
Center code: SOSTDC
Web site: http://sequence-www.stanford.edu/group/human/
Contact: hum-info@sequence.stanford.edu
----- Project Information
Center project name: 854
Center clone name: RP11-164C12
----- Summary Statistics
Sequencing Vector: M13mp18; X02513

```

```

Chemistry: Dye-primer; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 181410 bases at least Q40
Consensus quality: 181662 bases at least Q30
Consensus quality: 181772 bases at least Q20
Insert size: 182311; agarose-1p
Insert size: 182327; sum-of-contigs
Quality coverage: 10.2x in Q20 bases; agarose-1p
Quality coverage: 10.2x in Q20 bases; sum-of-contigs.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

```

```

* 1 17671: contig of 17671 bp in length
* 17672 17771: gap of unknown length
* 17772 55704: contig of 37933 bp in length
* 55705 55804: gap of unknown length
* 55805 114621: contig of 58817 bp in length
* 114622 114721: gap of unknown length
* 114722 182627: contig of 67906 bp in length.

```

FEATURES

source

```

1..182627
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="15"
/clone_lib="RP11-164C12"
/clone_lib="RP11-164C12"
/misc_feature 1..17671
/feature="assembly_name:Contig10"
17772..55704
/misc_feature /feature="assembly_name:Contig11"
55805..114621
/misc_feature /feature="assembly_name:Contig12"
114722..182627
/misc_feature /feature="assembly_name:Contig13"

```

```

BASE COUNT 53714 a 38503 c 38339 g 51766 t 305 others
ORIGIN

```

```

Query Match          6.6%; Score 66; DB 2; Length 182627;
Best Local Similarity 100.0%; Pred. NO. 3.9e-27;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 346 AAATACAAAATTAGCCAGGTGTGTGGCACAGCGCTAGTCCAGCTACTTGGGAGGC 405
|||||
Db 56133 AAATACAAAATTAGCCAGGTGTGTGGCACAGCGCTAGTCCAGCTACTTGGGAGGC 56192

```

```

QY 406 TGAGGC 411
|||||
Db 56193 TGAGGC 56198

```

RESULT 25

AC091078/c

LOCUS

AC091078 Homo sapiens, clone RP11-164C12, complete sequence.

AC091078

AC091078

AC091078.7 GI:21700642

KEYWORDS

HTG.

SOURCE

Homo sapiens.

Homo sapiens

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 182776)

Birren, B., Nusbaum, C. and Lander, E.

Homo sapiens, clone RP11-164C12

Unpublished

2 (bases 1 to 182776)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
 Barina,N., Bastien,V., Boguslavsky,L., Boukhgalter,B., Brown,A.,
 Camarata,J., Campopiano,A., Chang,J., Choepel,Y., Colangelo,M.,
 Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K.,
 Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D.,
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
 Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W.,
 Iliev,I., Johnson,R., Jones,C., Karatas,A., Larocque,K.,
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
 MacLean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M.,
 McWan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L.,
 Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,
 Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
 Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
 Retta,R., Riback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
 Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
 Severy,P., Sougnez,C., Spencer,B., Stange-Thomann,N.,
 Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
 Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE

Direct Submission

JOURNAL

Submitted (28-MAR-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 182776)

REFERENCE

Submitted (28-MAR-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 182776)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
 Barina,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
 Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
 Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
 Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
 Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
 McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
 Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
 Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
 Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
 Zembek,L., Zimmer,A. and Zody,M.

TITLE

Direct Submission

JOURNAL

Submitted (06-JUL-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 182776)

REFERENCE

Submitted (06-JUL-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 182776)

AUTHORS

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
 Barina,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
 Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
 Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
 Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
 Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
 McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
 Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
 Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
 Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
 Zembek,L., Zimmer,A. and Zody,M.

TITLE

Direct Submission

JOURNAL

Submitted (25-JUL-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jul 6, 2002 this sequence version replaced gi:21392507.

COMMENT

All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RW/RepeatMasker.html
 ----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WtBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: LI2380
 Center clone name: 164_C_12

FEATURES

source	Location/Qualifiers
repeat_region	1..154 /rpt_family="Charliel"
repeat_region	155..456 /rpt_family="AluY"
repeat_region	457..551 /rpt_family="Charliel"
repeat_region	590..929 /rpt_family="LIM4"
repeat_region	complement(930..1228) /rpt_family="AluSp"
repeat_region	1229..1284 /rpt_family="LIM4"
repeat_region	complement(1572..2004) /rpt_family="MLTIC"
repeat_region	2262..2287 /rpt_family="(T)n"
Repeat_region	2288..2458 /rpt_family="AluSg/x"
repeat_region	2450..2685 /rpt_family="LIMed"
repeat_region	2845..3116 /rpt_family="LIMed"
repeat_region	3139..3295 /rpt_family="LIMed"
repeat_region	complement(3296..3601) /rpt_family="AluSx"
repeat_region	3602..3662 /rpt_family="LIMed"
repeat_region	4030..4186 /rpt_family="LIMed"
repeat_region	4940..5021 /rpt_family="MIR"
repeat_region	5956..5965 /rpt_family="AT-rich"
repeat_region	complement(5966..6290) /rpt_family="LIPB1"
repeat_region	complement(6578..6693) /rpt_family="AluJb"
repeat_region	6816..6838 /rpt_family="(TTCA)n"
repeat_region	complement(7035..7086) /rpt_family="L2"
repeat_region	8170..8622 /rpt_family="MLTIC"
repeat_region	8715..8785 /rpt_family="MIR3"
repeat_region	9382..9588 /rpt_family="L3"
repeat_region	9940..9973 /rpt_family="(CA)n"
repeat_region	9974..10017 /rpt_family="(GA)n"
repeat_region	complement(10608..10909) /rpt_family="AluY"
repeat_region	complement(11914..11975) /rpt_family="MER5B"
repeat_region	11976..12279 /rpt_family="AluSp"
repeat_region	12288..12583

```

repeat_region      /rpt_family="AluJb"
                    complement(12584..12703)
repeat_region      /rpt_family="MERSB"
                    12714..12793
repeat_region      /rpt_family="MIR"
                    complement(13588..13770)
repeat_region      /rpt_family="LIMA5A"
                    13765..13838
repeat_region      /rpt_family="LIMA5A"
                    16006..16312
repeat_region      /rpt_family="AluY"
                    17503..17620
repeat_region      /rpt_family="MER46C"
                    18030..18330
repeat_region      /rpt_family="AluY"
                    complement(18370..18544)
repeat_region      /rpt_family="AluX"
                    complement(18545..18848)
repeat_region      /rpt_family="AluSg"
                    complement(18849..18981)
repeat_region      /rpt_family="AluX"
                    19509..19566
repeat_region      /rpt_family="MER46C"
                    complement(20146..20513)
repeat_region      /rpt_family="MIR1A1"
                    complement(20681..20719)
repeat_region      /rpt_family="AluSg"
                    complement(20720..20796)
repeat_region      /rpt_family="Alu"
                    21232..21316
repeat_region      /rpt_family="ORSL"
                    21325..21622
repeat_region      /rpt_family="AluJb"
                    21651..21757
repeat_region      /rpt_family="ORSL"
                    21994..22045
repeat_region      /rpt_family="AT_rich"
                    22163..22214
repeat_region      /rpt_family="T-rich"
                    complement(22321..28436)

```

Query Match 6.6%; Score 66; DB 9; Length 182776;

Best Local Similarity 100.0%; Pred. No. 3.9e-27;

Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 346 AAATACAAAATAGCCAGGTGGTGGCACACGCTGTAGTCCACGCTACTTGGGAGGC 405

Db 182452 AAATACAAAATAGCCAGGTGGTGGCACACGCTGTAGTCCACGCTACTTGGGAGGC 182393

QY 406 TGAGGC 411

Db 182392 TGAGGC 182387

RESULT 26
AC022281/c
LOCUS AC022281 186431 bp DNA linear HTG 24-JAN-2002
DEFINITION Homo sapiens chromosome 10 clone RP11-10E13; WORKING DRAFT
SEQUENCE, 9 unordered pieces.
ACCESSION AC022281
VERSION AC022281.3 GI:13399348
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP; HTGS_ACTIVEFIN;
HTGS_CANCELLED.
SOURCE Homo sapiens.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 186431)

Smith,D.R.

Genome Therapeutics Corporation Sequencing Center: Human Genome

Sequence Data

Unpublished

2 (bases 1 to 186431)

AUTHORS
TITLE
JOURNAL
COMMENT

Smith,D.R.
Direct Submission
Submitted (28-JAN-2000) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
On Mar 21, 2001 this sequence version replaced gi:9929643.

Genome Center
Center: Genome Therapeutics Corporation
Center code: GTC
Web site: http://www.genomecorp.com/
Contact: gtc-seqcenter@genomecorp.com

Project Information
Center project name: hgl43

Summary Statistics
Sequencing vector: N/A
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 990315
Consensus quality: 395389 bases at least Q40
Consensus quality: 405067 bases at least Q30
Consensus quality: 410525 bases at least Q20
Insert size: 185631; sum-of-contigs
Quality coverage: 6.4x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 124725: contig of 124725 bp in length
* 124726 124825: gap of unknown length
* 124826 125924: contig of 1099 bp in length
* 125925 126024: gap of unknown length
* 126025 128493: contig of 2469 bp in length
* 128494 128593: gap of unknown length
* 128594 156515: contig of 27922 bp in length
* 156516 156615: gap of unknown length
* 156616 170011: contig of 13396 bp in length
* 170012 170111: gap of unknown length
* 170112 176430: contig of 6319 bp in length
* 176431 176530: gap of unknown length
* 176531 179019: contig of 2489 bp in length
* 179020 179120: gap of unknown length
* 179120 181965: contig of 2846 bp in length
* 181966 182065: gap of unknown length
* 182066 186431: contig of 4366 bp in length.

FEATURES
source

1..186431
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-10E13"
/clone_lib="RPC1-11"
1..124725
/note="assembly_name:Contig243
clone_end:SP6"

misc_feature

124826..125924
/note="assembly_name:Contig204"
126025..128493
/note="assembly_name:Contig233"

misc_feature

128594..156515
/note="assembly_name:Contig240"
156616..170011
/note="assembly_name:Contig239"

misc_feature

170112..176430
/note="assembly_name:Contig235"
176531..179019
/note="assembly_name:Contig231"

misc_feature

179120..181965
/note="assembly_name:Contig232"
182066..186431
/note="assembly_name:Contig236"

```

BASE COUNT 48827 a 41798 c 41896 g 53062 t 848 others
ORIGIN
Query Match 6.6%; Score 66; DB 2; Length 186431;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 GTGGATCACTTGGTTCAGGAGTTTTCAGACACGCTGGCCACACGGTGAACCCCAT 201
|||||
Db 41809 GTGGATCACTTGGTTCAGGAGTTTTCAGACACGCTGGCCACACGGTGAACCCCAT 41750
|||||

QY 202 CTCCTAC 207
|||||
Db 41749 CTCCTAC 41744

RESULT 27
CNSOLDWM 189116 bp DNA linear PRI 08-JAN-2002
LOCUS Human chromosome 14 DNA sequence BAC C-2017C7 of library CalTech-D
DEFINITION from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION AL137779
VERSION AL137779.6 GI:18103850
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 189116)
AUTHORS Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,
Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C.,
Gyapay,G., Saurin,K. and Weissbach,J.
TITLE Sequencing of the human chromosome 14
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 189116)
Genoscope.
Direct Submission
Submitted (08-JAN-2002) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
On Jan 9, 2002 this sequence version replaced gi:15718346.
----- Genome Center
Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: SeqRef@genoscope.cns.fr
-----
The following BAC sequence is oriented from the T7 to the SP6 end.
Upstream BAC (overlapping the T7 end) : R-1017G21 (AC-ALL18558)
Downstream BAC (overlapping the SP6 end) : R-79666 -----
Summary Statistics
Assembly program: Phrap; version 2.0
Quality coverage: 7.29x in Q20 bases; sum-of-contigs
-----
Overall quality chart :
Range : bases
0 : 2
1 - 9 : 47
10 - 19 : 204
20 - 29 : 463
30 - 39 : 1585
40 - 49 : 5020
50 - 59 : 8017
60 - 69 : 8705
70 - 79 : 19573
80 - 89 : 53510
90 - 99 : 91990
-----
Percentage of bases with a quality value >= 40 : 98 %.
Location/Qualifiers

```

```

1. 189116
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="14"
/clone="C-2017C7"
/clone_lib="CalTech-D"
33932..34056
/notes="matching EMBL:T96367
RHdb:RH53659
dbSTS:STS4719
Identified using the e-PCR software (G. Schuler)"
42542..42731
/notes="matching EMBL:AA236880
RHdb:RH91926
dbSTS:STS64930
Identified using the e-PCR software (G. Schuler)"
80798..80947
/notes="matching EMBL:T88698
RHdb:RH53754
dbSTS:STS31753
Identified using the e-PCR software (G. Schuler)"
89772..89898
/notes="matching EMBL:H64944
RHdb:RH53858
dbSTS:STS24743
Identified using the e-PCR software (G. Schuler)"
101676..101841
/notes="matching EMBL:T52954
RHdb:RH44791
dbSTS:STS37859
Identified using the e-PCR software (G. Schuler)"
114961..115136
/notes="matching EMBL:Z51483
RHdb:RH31332
dbSTS:STS25252
Identified using the e-PCR software (G. Schuler)"

BASE COUNT 53622 a 42176 c 42426 g 50890 t 2 others
ORIGIN
Query Match 6.6%; Score 66; DB 9; Length 189116;
Best Local Similarity 100.0%; Pred. No. 3.9e-27;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 GTGGATCACTTGGTTCAGGAGTTTTCAGACACGCTGGCCACACGGTGAACCCCAT 201
|||||
Db 45000 GTGGATCACTTGGTTCAGGAGTTTTCAGACACGCTGGCCACACGGTGAACCCCAT 45059
|||||

QY 202 CTCCTAC 207
|||||
Db 45060 CTCCTAC 45065

RESULT 28
AL136137 127140 bp DNA linear PRI 06-DEC-2001
LOCUS Human DNA sequence from clone RP11-377B2 on chromosome Xp11.23-11.4
DEFINITION Contains a novel pseudogene, a pseudogene similar to RPL19
(ribosomal protein L19), complete sequence.
ACCESSION AL136137
VERSION AL136137.15 GI:10119664
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 127140)
AUTHORS Whitehead,S.
TITLE Direct Submission
JOURNAL Submitted (06-DEC-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Sep 12, 2000 this sequence version replaced gi:10086012.
During sequence assembly data is compared from overlapping clones.
COMMENT

```

Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/Projects/C-elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/ChrX>

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP11-377B2 is from the library RPCI-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-377B2. It may be shorter because we sequence overlapping sections only once, except for a short overlap.

The true left end of clone RP6-105D16 is at 127041 in this sequence. The true right end of clone RP4-551E13 is at 100 in this sequence.

FEATURES

source	Location/Qualifiers
misc_feature	1..127140
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="X"
	/map="p11.23-11.4"
	/clone_lib="RPCI-11.2"
	/clone="RP11-377B2"
	complement(1..428)
	/note="match: GSS: Em:AQ027029"
	972..1271
repeat_region	/note="AluY repeat: matches 2..302 of consensus"
misc_feature	complement(1321..1598)
	/note="match: STS: Em:L47514"
	2140..2253
repeat_region	/note="FLAM_C repeat: matches 1..117 of consensus"
repeat_region	2602..2887
	/note="AluSx repeat: matches 1..286 of consensus"
	2912..3224
	/note="AluSx repeat: matches 1..303 of consensus"
	3238..3526
	/note="AluJo repeat: matches 1..290 of consensus"
	3527..3648
	/note="61 copies 2 mer ca 77% conserved"
	3830..3982
	/note="HAL1 repeat: matches 24..164 of consensus"
repeat_region	4044..4081
repeat_region	/note="19 copies 2 mer ac 100% conserved"
	4130..4390
	/note="L1MC3 repeat: matches 6875..7167 of consensus"
	4391..4691
	/note="AluSx repeat: matches 1..302 of consensus"
	4796..5146
	/note="match: GSS: Em:AQ633427"
	complement(4822..4963)
	/note="match: STS: Em:HSA058XH1"
	4836..5129
repeat_region	/note="AluSg repeat: matches 1..292 of consensus"
repeat_region	5626..5935
	/note="AluY repeat: matches 1..311 of consensus"
	6060..6377
	/note="AluJb repeat: matches 1..312 of consensus"
	6384..6444
	/note="HAL1 repeat: matches 289..354 of consensus"
	6492..6502
	/note="BC200 repeat: matches 1..10 of consensus"
	6503..6621
	/note="AluJo repeat: matches 10..139 of consensus"
repeat_region	6622..6641
	/note="BC200 repeat: matches 10..161 of consensus"
	6642..7142
	/note="HAL1 repeat: matches 357..947 of consensus"
	7143..7443
	/note="AluSx repeat: matches 1..305 of consensus"
	7444..7555
	/note="HAL1 repeat: matches 947..1044 of consensus"
	7664..7977
	/note="AluY repeat: matches 1..311 of consensus"
repeat_region	7982..8294
	/note="AluSg repeat: matches 1..309 of consensus"
	8302..8615
	/note="AluSg repeat: matches 1..299 of consensus"
	8632..8733
	/note="L1MAC repeat: matches 1564..1666 of consensus"
	8861..9122
	/note="HAL1 repeat: matches 1427..1704 of consensus"
	10252..10559
	/note="AluSg repeat: matches 3..308 of consensus"
repeat_region	11009..11135
	/note="L1MC5 repeat: matches 7754..7900 of consensus"
	11573..11941
	/note="MUTIC repeat: matches 6..416 of consensus"
	11955..12244
	/note="AluSx repeat: matches 1..291 of consensus"
	12693..12724
	/note="16 copies 2 mer tg 96% conserved"
	15518..15901
	/note="MER31A repeat: matches 1..375 of consensus"
repeat_region	MER31A repeat: matches 1..375 of consensus"
	15902..16194
	/note="AluJo repeat: matches 1..292 of consensus"
	16195..16277
	/note="MER31A repeat: matches 375..458 of consensus"
	MER31A repeat: matches 375..458 of consensus"
	16586..16893
	/note="AluSx repeat: matches 1..301 of consensus"
	17240..17283
	/note="MADEL repeat: matches 35..80 of consensus"
repeat_region	17638..18226
	/note="LTR10C repeat: matches 1..586 of consensus"
	LTR10C repeat: matches 1..586 of consensus"
	18309..18597
	/note="AluSp repeat: matches 1..288 of consensus"
	18621..18909
	/note="AluSx repeat: matches 1..284 of consensus"
	19245..19569
	/note="match: GSS: Em:AQ663911 Em:B94529"
	19245..19512
misc_feature	/note="match: GSS: Em:AQ128461"
	complement(19373..19569)
	/note="match: GSS: Em:AQ240745"
	19381..19521
	/note="match: GSS: Em:AQ037473"
	19405..19614
	/note="MER4D repeat: matches 764..981 of consensus"
	19615..19829
	/note="AluSg/X repeat: matches 92..306 of consensus"
	19830..20397
repeat_region	/note="MER4D repeat: matches 233..764 of consensus"
	20430..20567
	/note="AluJo repeat: matches 1..119 of consensus"
	20568..20869
	/note="AluSx repeat: matches 1..301 of consensus"
	20870..20999
	/note="AluJo repeat: matches 119..306 of consensus"
	20999..21111
	/note="AluY repeat: matches 1..311 of consensus"
	21111..21222

```

repeat_region 21447..21599
/Note="LTR8 repeat: matches 1..156 of consensus
LTR8 repeat: matches 1..156 of consensus"
21600..21919
/Note="AluX repeat: matches 1..310 of consensus"
21920..22306
/Note="LTR8 repeat: matches 156..524 of consensus
LTR8 repeat: matches 156..524 of consensus"
22307..22614
/Note="AluS8 repeat: matches 1..310 of consensus"
22615..22779
/Note="LTR8 repeat: matches 524..691 of consensus
LTR8 repeat: matches 524..691 of consensus"
22930..23226
/Note="AluS8 repeat: matches 1..310 of consensus"
23239..23324
/Note="LTR8 repeat: matches 6..302 of consensus"
23339..23324
/Note="MER57-internal repeat: matches 7149..7239 of
consensus"
23449..23755
/Note="match: GSS: Em:AQ826300"
23464..23755
/Note="AluX repeat: matches 3..311 of consensus"
24354..24509
/Note="match: STS: Em:HS138B7S"
24510..24806
/Note="AluY repeat: matches 137..292 of consensus"
24807..24943
/Note="AluY repeat: matches 1..297 of consensus"
25494..25683
/Note="AluY repeat: matches 1..137 of consensus"
25494..25683
/Note="MER58C repeat: matches 1..81 of consensus
MER58C repeat: matches 1..81 of consensus"
25685..25976
/Note="match: GSS: Em:AQ590309"

Query Match 6.4%; Score 64; DB 9; Length 127140;
Best Local Similarity 100.0%; Pred. No. 5.8e-26;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 CGTCTCTACTAAATAACAAAATAGCCAGGTGGTGGGCACACGCCCTGTAGTCCAGC 393
Db 51506 CGTCTCTACTAAATAACAAAATAGCCAGGTGGTGGGCACACGCCCTGTAGTCCAGC 51565

QY 394 TACT 397
|||||
Db 51566 TACT 51569

RESULT 29
AL590640
LOCUS
DEFINITION Human DNA sequence from clone RP11-40H20 on chromosome 1, complete
sequence.
ACCESSION AL590640
VERSION AL590640.18 GI:18476678
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 146954)
Harrison,E.
Direct Submission
Submitted (30-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk
On Feb 1, 2002 this sequence version replaced gi:18181688.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with

```

```

only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sg:
SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chrl
RP11-40H20 is from the library RPI-11.1 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-40H20 it may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true left end of clone RP11-40H20 is at 1 in this sequence. The
true left end of clone RP4-633N17 is at 144955 in this sequence.
The true right end of clone RP11-4K3 is at 7274 in this sequence.
LOCATION/Qualifiers
source 1..146954
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-40H20"
/clone_id="RPI-11.1"
5405..5468
misc_feature
/Note="Sequence from overlapping clone RP11-4K3
(AL591050). Assembly confirmed by restriction digest."
26429..26442
misc_feature
/Note="Single clone region. Reads generated from a
transposon library derived from a single pUC clone.
Restriction digest data confirm the assembly."
108250..108261
misc_feature
/Note="Sequence from uni-directional dGTP big dye
terminator reads only."
Single clone region. Reads generated from a transposon
library derived from a single pUC clone. Restriction
digest data confirm the assembly."
139312..139369
misc_feature
/Note="Sequence from uni-directional dGTP big dye
terminator reads only."
BASE COUNT 41970 a 33288 c 33765 g 37930 t
ORIGIN

Query Match 6.4%; Score 64; DB 9; Length 146954;
Best Local Similarity 100.0%; Pred. No. 5.9e-26;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 CGTCTCTACTAAATAACAAAATAGCCAGGTGGTGGGCACACGCCCTGTAGTCCAGC 393
|||||
Db 40427 CGTCTCTACTAAATAACAAAATAGCCAGGTGGTGGGCACACGCCCTGTAGTCCAGC 40486

QY 394 TACT 397
|||||
Db 40487 TACT 40490

RESULT 30
AL359272
LOCUS
DEFINITION Human DNA sequence from clone RP11-554P16 on chromosome X. Contains
the first coding exon of the gene KIAA1202, STSs, GSSs and a CpG
island, complete sequence.
ACCESSION AL359272

```

VERSION
KEYWORDS
SOURCE

AL359272.9 GI:9650579
HTG: Cpg island; KIAA1202.
SOURCE human.

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 173510)

AUTHORS

Bird,C.

JOURNAL

Direct Submission

COMMENT

Submitted (28-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Aug 1, 2000 this sequence version replaced gi:9588592.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
<http://www.sanger.ac.uk/Projects/C.elegans/wormpep> This sequence
was generated from part of bacterial clone contigs of human
chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at
<http://www.sanger.ac.uk/RGP/chrX>

RP11-554P16 is from the library RPCI-11.2 constructed at the
Roswell Park Cancer Institute by the group of Pieter de Jong. For
further details see <http://bacpac.med.buffalo.edu/>
VECTOR: pBACe3.6

This sequence is the entire insert of clone RP11-554P16.

FEATURES

source

1. .173510
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP11-554P16"
/clone_lib="RPCI-11.2"

repeat_region

/note="LM3C repeat: matches 1913. .2098 of consensus"
7. .674

misc_feature

/note="match: GSS: Em:AQ373815"

repeat_region

195. .511
/note="AluSp repeat: matches 1. .312 of consensus"

repeat_region

543. .845
/note="AluXs repeat: matches 1. .304 of consensus"

repeat_region

950. .1127
/note="LM3C repeat: matches 1968. .2136 of consensus"

repeat_region

1190. .2306
/note="LM4b repeat: matches -180. .888 of consensus"

repeat_region

2425. .3343
/note="LM4 repeat: matches 3078. .4092 of consensus"

repeat_region

3344. .3717
/note="MSTD repeat: matches 1. .394 of consensus"

repeat_region

3718. .4412
/note="LM4 repeat: matches 4092. .4274 of consensus"

repeat_region

4413. .4452
/note="THE1B repeat: matches 1. .42 of consensus"

repeat_region

4453. .4763
/note="AluXs repeat: matches 1. .310 of consensus"

repeat_region

4764. .5111
/note="THE1B repeat: matches 42. .364 of consensus"

repeat_region

5112. .5774
/note="LM4 repeat: matches 4274. .5344 of consensus"

repeat_region

5775. .6080

repeat_region

/note="AluSg repeat: matches 1. .307 of consensus"
6081. .6296

repeat_region

/note="LM4 repeat: matches 5344. .5581 of consensus"
6299. .6923

repeat_region

/note="LM10 repeat: matches 5525. .6165 of consensus"
7394. .7652

repeat_region

/note="LM4 repeat: matches 2120. .2383 of consensus"
7652. .10378

repeat_region

/note="LM4 repeat: matches 3414. .6142 of consensus"
10369. .10878

repeat_region

/note="LM4 repeat: matches 2350. .2879 of consensus"
10875. .11250

repeat_region

/note="L1 repeat: matches 4741. .5131 of consensus"
11247. .11619

repeat_region

/note="LM4 repeat: matches 5903. .6284 of consensus"
11642. .11764

repeat_region

/note="LM4 repeat: matches 5351. .5475 of consensus"
11764. .11943

repeat_region

/note="L1 repeat: matches 3846. .4036 of consensus"
12269. .13359

repeat_region

/note="LM3 repeat: matches 6659. .7739 of consensus"
complement(13126. .13490)

repeat_region

/note="match: STS: Em:HS308ZG9"
13430. .13497

repeat_region

/note="LM4 repeat: matches 1995. .2062 of consensus"
13498. .13803

repeat_region

/note="AluJo repeat: matches 3. .307 of consensus"
13804. .15085

repeat_region

/note="LM4 repeat: matches 720. .1995 of consensus"
15191. .15376

repeat_region

/note="LM4 repeat: matches 408. .602 of consensus"
15565. .16035

repeat_region

/note="LM4 repeat: matches -252. .225 of consensus"
16038. .16114

repeat_region

/note="L2 repeat: matches 2686. .2745 of consensus"
16128. .16369

repeat_region

/note="LM4 repeat: matches 5928. .6318 of consensus"
17285. .17533

repeat_region

/note="match: GSS: Em:AQ134667"
17455. .17899

repeat_region

/note="match: GSS: Em:AQ142071"
17875. .18018

repeat_region

/note="MIR repeat: matches 96. .252 of consensus"
18979. .19286

repeat_region

/note="AluJb repeat: matches 2. .308 of consensus"
19391. .19823

repeat_region

/note="L2 repeat: matches 2284. .2708 of consensus"
20320. .20610

repeat_region

/note="AluX repeat: matches 1. .289 of consensus"
21916. .22121

repeat_region

/note="match: STS: Em:G03829"
23553. .23862

repeat_region

/note="MER33 repeat: matches 1. .307 of consensus"
23946. .24118

repeat_region

/note="MIR repeat: matches 77. .259 of consensus"
24356. .24475

repeat_region

/note="L2 repeat: matches 2624. .2750 of consensus"
24929. .25125

repeat_region

/note="MIR repeat: matches 23. .256 of consensus"
25251. .25386

repeat_region

/note="MER63 repeat: matches 2. .768 of consensus"
25604. .25656

repeat_region

/note="L2 repeat: matches 2657. .2709 of consensus"
26047. .27589

repeat_region

/note="LM8 repeat: matches 4591. .6159 of consensus"
27590. .27886

repeat_region

/note="AluX repeat: matches 1. .297 of consensus"
27887. .28214

repeat_region

/note="LM8 repeat: matches 4271. .4591 of consensus"
28258. .29096

repeat_region

/note="LM8 repeat: matches 3264. .4157 of consensus"
29118. .29521

repeat_region

/note="LM1B repeat: matches 1. .383 of consensus"

repeat_region 29523. .30381
/note="L1MB8 repeat: matches 2496. .3231 of consensus"
repeat_region 30382. 30685
/note="AluJo repeat: matches 2. .299 of consensus"
repeat_region 30686. .31069
/note="L1MB8 repeat: matches 2110. .2496 of consensus"
repeat_region 31088. .31421
/note="L1PA13 repeat: matches 5820. .6156 of consensus"
repeat_region 31422. .31617
/note="L1P4 repeat: matches 5374. .5552 of consensus"
repeat_region 31618. .31909
/note="AluX repeat: matches 1. .294 of consensus"
repeat_region 31910. .31976
/note="L1P4 repeat: matches 5304. .5374 of consensus"
repeat_region 31978. .32359
/note="L1MC repeat: matches 1951. .2330 of consensus"
repeat_region 32354. .33592
/note="L1M4C repeat: matches 969. .2197 of consensus"
repeat_region 33935. .34790
/note="L1M4B repeat: matches -266. .602 of consensus"
repeat_region 34856. .35219
/note="THE1B repeat: matches 1. .364 of consensus"
repeat_region 36043. .36159
/note="L2 repeat: matches 2576. .2710 of consensus"
misc_feature 38455. .39057
/note="match: GSS: Em:AQ008362"
misc_feature 41978. .42373
/note="match: STS: Em:L18363"
repeat_region 42056. .42087
/note="16 copies 2 mer tg 96% conserved"
repeat_region 42311. .42540
/note="MIR repeat: matches 19. .262 of consensus"
repeat_region 42824. .42914
/note="L2 repeat: matches 2606. .2709 of consensus"
repeat_region 43119. .43173
/note="L2 repeat: matches 2687. .2742 of consensus"
repeat_region 43122. .43185
/note="MIR repeat: matches 194. .257 of consensus"
repeat_region 44700. .44743
/note="22 copies 2 mer aa 75% conserved"
repeat_region 45956. .46137
/note="MIR repeat: matches 20. .199 of consensus"
repeat_region 46401. .46617
/note="MIR repeat: matches 23. .262 of consensus"
repeat_region 46729. .47035
/note="AluJb repeat: matches 1. .309 of consensus"
repeat_region 47439. .47682
/note="L1ME1 repeat: matches 5666. .5914 of consensus"
repeat_region 48430. .48527

Query Match 6.4%; Score 64; DB 9; Length 173510;
Best Local Similarity 100.0%; Pred. No. 5.9e-26;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 CGTCTCTACTAAATAACAAAATAGCCAGGTGGTGGCACACGCCCTGTAGTCCGACG 393
Db 169132 CCGTCTCTACTAAATAACAAAATAGCCAGGTGGTGGCACACGCCCTGTAGTCCGACG 169191

QY 394 TACT 397
|||||
Db 169192 TACT 169195

RESULT 31
AC021923
LOCUS AC021923 177242 bp DNA linear HTG 09-MAY-2001
DEFINITION Homo sapiens clone RP11-2707, WORKING DRAFT SEQUENCE, 7 unordered
pieces.
ACCESSION AC021923
VERSION AC021923.4 GI:12656831
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens.
ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 177242)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Unpublished
2 (bases 1 to 177242)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,H.,
Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G., Castle,A.,
Choeple,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArelano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferrela,P., FitzHugh,W., Forrest,C., Gage,B., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lechovsky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
MacDonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheders,R., Meldrum,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Ollivar,T.M., Peterson,K.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 2, 2001 this sequence version replaced gi:7249099.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L4730
Center clone name: 27_0_7
----- Summary Statistics
Sequencing vector: M13; M77815; 49% of reads
Sequencing vector: Plasmid; n/a; 51% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 175551 bases at least Q40
Consensus quality: 176231 bases at least Q30
Consensus quality: 176426 bases at least Q20
Insert size: 180000; agarose-fp
Insert size: 176542; sum-of-contigs
Quality coverage: 9.8 in Q20 bases; agarose-fp
Quality coverage: 10.0 in Q20 ba.

TITLE
JOURNAL
COMMENT

* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 3850: contig of 3850 bp in length
* 3851 3950: gap of 100 bp
* 3951 9386: contig of 5436 bp in length
* 9387 9486: gap of 100 bp
* 9487 17804: contig of 8318 bp in length
* 17805 17904: gap of 100 bp
* 17905 71191: contig of 53287 bp in length
* 71192 71291: gap of 100 bp
* 71292 112250: contig of 40959 bp in length
* 112251 112350: gap of 100 bp
* 112351 154709: contig of 42359 bp in length
* 154710 154809: gap of 100 bp
* 154810 177242: contig of 22433 bp in length.
* Location/Qualifiers

FEATURES

```

source
1. .177242
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-2707"
/clone_lib="RPC1-11 Human Male BAC"
1. .3850
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
misc_feature
3951. .9386
/note="assembly_fragment"
9487. .17804
/note="assembly_fragment"
17905. .71191
/note="assembly_fragment"
71292. .112250
/note="assembly_fragment"
112351. .154709
/note="assembly_fragment"
154810. .177242
/note="assembly_fragment"
clone_end:T7
vector_side:right"
BASE COUNT 49940 a 37600 c 38021 g 51078 t 603 others
ORIGIN

Query Match 6.4%; Score 64; DB 2; Length 177242;
Best Local Similarity 100.0%; Pred. No. 5.9e-26;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GGGTGGCGGTGCACACCTGTAATCCAGCACCTTGGAGGCTGAGACGGGTGGATCAC 287
|||||
Db 85852 GGGTGGCGGTGCACACCTGTAATCCAGCACCTTGGAGGCTGAGACGGGTGGATCAC 85911
|||||

QY 208 CTGA 291
|||||
Db 85912 CTGA 85915

RESULT 32
AP003352/c
LOCUS 187349 bp DNA linear PRI 03-APR-2001
DEFINITION Homo sapiens genomic DNA, Chromosome 8q23, clone: KB1208A12.
ACCESSION AP003352
VERSION AP003352.2 GI:13516385
KEYWORDS
SOURCE Homo sapiens cell_line:FLEB 14 - 14 DNA, clone_lib:Keio BAC library
clone:KB1208A12.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. (sites)
Shimizu,N. and Asakawa,S.
TITLE Homo sapiens DNA chromosome 8 SEQUENCE
JOURNAL Published Only in Database (2001)
REFERENCE 2 (bases 1 to 187349)
Shimizu,N. and Asakawa,S.
TITLE Direct Submission
JOURNAL Submitted (28-FEB-2001) Nobuyoshi Shimizu, Keio University, School
of Medicine, Molecular Biology, 35 Shinanomachi, Shinjuku-ku, Tokyo
160-8582, Japan (E-mail:ashimizu@cmb.med.keio.ac.jp,
Tel:81-3-3351-2370, Fax:81-3-3351-2370)
COMMENT On Apr 2, 2001 this sequence version replaced gi:13359068.
FEATURES
Location/Qualifiers
1. .187349
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8q23"
/clone="KB1208A12"
/cell_line="FLEB 14 - 14"
/clone_lib="Keio BAC library"
repeat_region 307. .407

```

```

/evidence-not_experimental
/rpt_family="L2"
complement(752. .1291)
/evidence-not_experimental
/rpt_family="LIM4"
1735. .2095
/evidence-not_experimental
/rpt_family="Charlie4a"
complement(2109. .2538)
/evidence-not_experimental
/rpt_family="LIMB5"
complement(2539. .2849)
/evidence-not_experimental
/rpt_family="AluSg"
complement(2850. .3129)
/evidence-not_experimental
/rpt_family="LIMB5"
3130. .3228
/evidence-not_experimental
/rpt_family="Charlie4a"
complement(3887. .4189)
/evidence-not_experimental
/rpt_family="AluSx"
complement(4202. .4509)
/evidence-not_experimental
/rpt_family="AluY"
4526. .4729
/evidence-not_experimental
/rpt_family="MIR"
4731. .4752
/evidence-not_experimental
/rpt_family="AT-rich"
4762. .4925
/evidence-not_experimental
/rpt_family="AluSg"
5042. .5157
/evidence-not_experimental
/rpt_family="AluSx"
5217. .5339
/evidence-not_experimental
/rpt_family="FLAM_C"
complement(5513. .5813)
/evidence-not_experimental
/rpt_family="AluSx"
complement(5835. .6120)
/evidence-not_experimental
/rpt_family="AluJb"
complement(6185. .6288)
/evidence-not_experimental
/rpt_family="LIMB8"
6308. .6337
/evidence-not_experimental
/rpt_family="(TTTC)n"
complement(6374. .6674)
/evidence-not_experimental
/rpt_family="L2"
complement(6675. .6951)
/evidence-not_experimental
/rpt_family="AluY"
complement(6954. .7245)
/evidence-not_experimental
/rpt_family="AluSx"
complement(7246. .7360)
/evidence-not_experimental
/rpt_family="L2"
7570. .7814
/evidence-not_experimental
/rpt_family="MIR"
complement(7870. .8167)
/evidence-not_experimental
/rpt_family="AluSx"
8317. .8619
/evidence-not_experimental

```

```

repeat_region /rpt_family="AluSc"
complement(8629..8716)
/evidence-not_experimental
repeat_region /rpt_family="MIR"
complement(8754..8929)
/evidence-not_experimental
repeat_region /rpt_family="MIR"
complement(10203..10502)
/evidence-not_experimental
repeat_region /rpt_family="AluSx"
complement(10504..10802)
/evidence-not_experimental
repeat_region /rpt_family="AluSg"
complement(11247..11541)
/evidence-not_experimental
repeat_region /rpt_family="AluJb"
11669..11862
/evidence-not_experimental
repeat_region /rpt_family="L2"
complement(11863..12179)
/evidence-not_experimental
repeat_region /rpt_family="AluJo"
12180..12285
/evidence-not_experimental
repeat_region /rpt_family="L2"
12747..12781
/evidence-not_experimental
repeat_region /rpt_family="(TG)n"
12930..12973
/evidence-not_experimental
repeat_region /rpt_family="(TTTA)n"
complement(12974..13257)
/evidence-not_experimental
repeat_region /rpt_family="AluSx"
complement(13274..13504)
/evidence-not_experimental
repeat_region /rpt_family="L1PAl3"
13820..13844
/evidence-not_experimental
repeat_region /rpt_family="(TTTCC)n"
14691..14990
/evidence-not_experimental
repeat_region /rpt_family="AluSx"
15114..15249
/evidence-not_experimental
repeat_region /rpt_family="L2"
15666..15819
/evidence-not_experimental
repeat_region /rpt_family="MIR"
complement(16065..16162)
/evidence-not_experimental
repeat_region /rpt_family="MERSA"
complement(16166..16419)
/evidence-not_experimental
repeat_region /rpt_family="AluSg"
17087..17114
/evidence-not_experimental
repeat_region /rpt_family="(TAG)n"
complement(17115..17395)
/evidence-not_experimental
repeat_region /rpt_family="AluSx"
17570..17647
/evidence-not_experimental
repeat_region /rpt_family="AluSc"
18742..19066
/evidence-not_experimental
repeat_region /rpt_family="MER1B"
20245..20422
/evidence-not_experimental
repeat_region /rpt_family="L2"
20424..20452
/evidence-not_experimental
/rpt_family="(TC)n"

```

```

repeat_region complement(20453..20743)
/evidence-not_experimental
/rpt_family="AluSg"
20744..21255
/evidence-not_experimental
repeat_region /rpt_family="L2"
21787..21863
/evidence-not_experimental
/rpt_family="MIR"
complement(21864..22286)
/evidence-not_experimental
repeat_region /rpt_family="MER57A"
22287..22492
/evidence-not_experimental
/rpt_family="MIR"
complement(22461..22637)
/evidence-not_experimental
repeat_region /rpt_family="MIR"
23449..23484
/evidence-not_experimental
/rpt_family="MIR"
23897..24018

Query Match 6.4% Score 64; DB 9; Length 187349;
Best Local Similarity 100.0%; Pred. No. 6e-26;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GGGTGGGTGGCTCACACCTGTATCCAGCAGCTTTGGAGGCTGAGACGGTGGATCAC 287
|||||
Db 181566 GGGTGGGTGGCTCACACCTGTATCCAGCAGCTTTGGAGGCTGAGACGGTGGATCAC 181507

QY 288 CTGA 291
|||||
Db 181506 CTGA 181503

```

RESULT 33

AC091075

LOCUS

DEFINITION

AC091075

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AC091075 197870 bp DNA linear HTG 11-JUN-2001
Homo sapiens chromosome 8 clone RP11-680G2 map 8, WORKING DRAFT
SEQUENCE, 10 unordered pieces.
AC091075 2 GI:14336474
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 197870)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 8, clone RP11-680G2
Unpublished
2 (bases 1 to 197870)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Boguslavskiy,L., Bouckgeater,B., Brown,A.,
Cammarata,J., Campopiano,A., Chang,J., Choquet,Y., Colangelo,M.,
Collins,S., Collymore,A., Cooke,P., DeArelano,K., Dewar,K.,
Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W.,
Iliev,I., Johnson,R., Jones,C., Karatas,A., Lander,E., Lander,E.,
Lamarez,R., Landers,T., Lenocky,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., McPheeters,R., Meldrum,J., Meneus,L.,
Mihova,T., Mlekena,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,
Severy,P., Sougnez,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,

TITLE
JOURNAL
COMMENT

Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (28-MAR-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 11, 2001 this sequence version replaced gi:13470191.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L11753
Center clone name: 680_G_2
----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 194427 bases at least Q40
Consensus quality: 195744 bases at least Q30
Consensus quality: 196367 bases at least Q20
Insert size: 200000; agarose-fp
Quality coverage: 10.7 in Q20 bases; sum-of-contigs
Quality coverage: 10.9 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 36504: contig of 36504 bp in length
* 36505 36604: gap of 100 bp
* 36605 37316: contig of 712 bp in length
* 37317 37416: gap of 100 bp
* 37417 88008: contig of 50592 bp in length
* 88009 88108: gap of 100 bp
* 88109 89700: contig of 1592 bp in length
* 89701 89800: gap of 100 bp
* 89801 91377: contig of 1577 bp in length
* 91378 91477: gap of 100 bp
* 91478 96729: contig of 5252 bp in length
* 96730 96829: gap of 100 bp
* 96830 118877: contig of 22048 bp in length
* 118878 118977: gap of 100 bp
* 118978 146670: contig of 27693 bp in length
* 146671 146770: gap of 100 bp
* 146771 197149: contig of 50379 bp in length
* 197150 197249: gap of 100 bp
* 197250 197870: contig of 621 bp in length.

FEATURES

source
1. .197870
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone_lib="RP11-680G2"
/clone_lib="RP11-680G2"
1. .36504
/note="assembly_fragment
clone_end:SP6
vector_side:left"
36605..37316
/note="assembly_fragment"
37417..88008
/note="assembly_fragment"
88109..89700
/note="assembly_fragment"

misc_feature
89801..91377
/note="assembly_fragment"
91478..96729
misc_feature
/note="assembly_fragment"
96830..118877
misc_feature
/note="assembly_fragment"
118978..146670
misc_feature
/note="assembly_fragment"
146771..197149
misc_feature
/note="assembly_fragment"
197250..197870
misc_feature
/note="assembly_fragment
clone_end:T7
vector_side:right"
BASE COUNT 55077 a 43169 c 43473 g 55250 t 901 others
ORIGIN

Query Match 6.4%; Score 64; DB 2; Length 197870;
Best Local Similarity 100.0%; Pred.No. 6e-26;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 228 GGTCGGGTGGCTCACACCTGTATCCAGCACTTTGGAGGCTGAGACGGTGGATCAC 287
|||||
Db 9986 GGTCGGGTGGCTCACACCTGTATCCAGCACTTTGGAGGCTGAGACGGTGGATCAC 10045
QY 288 CTGA 291
|||||
Db 10046 CTGA 10049

RESULT 34
AC019236
LOCUS
DEFINITION Homo sapiens chromosome 8 clone RP11-680G2, WORKING DRAFT SEQUENCE,
23 unordered pieces.
AC019236 199208 bp DNA linear HTG 10-SEP-2000
VERSION AC019236.5 GI:10048101
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Waterston,R.H.
JOURNAL The sequence of Homo sapiens clone
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 199208)
TITLE Waterston,R.H.
JOURNAL Direct Submission
Submitted (30-DEC-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Sep 10, 2000 this sequence version replaced gi:7235318.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
----- Project Information -----
Center project name: HLNH0680G02
----- Summary Statistics -----
Sequencing vector: M13; 76%
Sequencing vector: plasmid; 24%
Chemistry: Dye-primer ET; 76% of reads
Assembly program: Big Dye; 24% of reads
Consensus quality: 188489 bases at least Q40
Consensus quality: 191861 bases at least Q30
Insert size: 202000; agarose-fp
Insert size: 198337; sum-of-contigs
Quality coverage: 4.00 in Q20 bases; agarose-fp
Quality coverage: 4.08 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 23 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1
1812: contig of 1812 bp in length
1813 1912: gap of unknown length
1913 3838: contig of 1926 bp in length
3839 3938: gap of unknown length
3939 5889: contig of 1951 bp in length
5890 5989: gap of unknown length
5990 8667: contig of 2678 bp in length
8668 8767: gap of unknown length
8768 11541: contig of 2774 bp in length
11542 11641: gap of unknown length
11643 14813: contig of 3172 bp in length
14814 14913: gap of unknown length
14914 17378: contig of 2465 bp in length
17379 17478: gap of unknown length
17479 21197: contig of 3719 bp in length
21198 21297: gap of unknown length
21298 26511: contig of 5214 bp in length
26512 26611: gap of unknown length
26612 30593: contig of 3984 bp in length
30594 30695: gap of unknown length
30696 35841: contig of 5146 bp in length
35842 35941: gap of unknown length
35942 39400: contig of 3459 bp in length
39401 39500: gap of unknown length
39501 43746: contig of 4246 bp in length
43747 43846: gap of unknown length
43847 49935: contig of 6089 bp in length
49936 50035: gap of unknown length
50036 54892: contig of 4857 bp in length
54893 54993: contig of 7611 bp in length
54994 62603: contig of 7611 bp in length
62604 62703: gap of unknown length
62704 70543: contig of 7840 bp in length
70544 70643: gap of unknown length
70644 79145: contig of 8502 bp in length
79146 79245: gap of unknown length
79246 88619: contig of 9374 bp in length
88620 88719: gap of unknown length
88720 100155: contig of 11436 bp in length
100156 100255: gap of unknown length
100256 114403: contig of 14148 bp in length
114404 114503: gap of unknown length
114504 142767: contig of 28264 bp in length
142768 142867: gap of unknown length
142868 199208: contig of 56341 bp in length.

FEATURES

source
1. .199208
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/clone="RP11-680G2"
1. .1812
/note="assembly_name:Contig4"
1913. .3838
/note="assembly_name:Contig5"
3939. .5889
/note="assembly_name:Contig6"
5990. .8667
/note="assembly_name:Contig7"
8768. .11541
/note="assembly_name:Contig8"
11642. .14813
/note="assembly_name:Contig9"
14914. .17378

misc_feature
/note="assembly_name:Contig10"
17479. .21197
/note="assembly_name:Contig11"
21298. .26511
/note="assembly_name:Contig12"
26612. .30595
/note="assembly_name:Contig13"
30696. .35841
/note="assembly_name:Contig14"
35942. .39400
/note="assembly_name:Contig15"
39501. .43746
/note="assembly_name:Contig16"
43847. .49935
/note="assembly_name:Contig17"
50036. .54892
/note="assembly_name:Contig18"
54993. .62603
/note="assembly_name:Contig19"
62704. .70543
/note="assembly_name:Contig20"
70644. .79145
/note="assembly_name:Contig21"
79246. .88619
/note="assembly_name:Contig22"
88720. .100155
/note="assembly_name:Contig23"
100256. .114403
/note="assembly_name:Contig24"
114504. .142767
/note="assembly_name:Contig25
clone_end:SP6
vector_side:left"
142868. .199208
/note="assembly_name:Contig26"
53771 a 43806 c 43167 g 56261 t 2203 others

Query Match 6.4%; Score 64; DB 2; Length 199208;
Best Local Similarity 100.0%; Pred. No. 6e-26;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GGGTGGCGTGGCTCACACCTGTATCCAGCAGCTTGGAGGCTGAGACGGTGGATCAC 287
|||||
Db 124488 GGGTGGCGTGGCTCACACCTGTATCCAGCAGCTTGGAGGCTGAGACGGTGGATCAC 124547

QY 288 CTGA 291
|||||
Db 124548 CTGA 124551

RESULT 35

LOCUS AL136973/c

DEFINITION Human DNA sequence from clone RP3-41408 on chromosome 6 Contains ESTs, STSS and GSSs, complete sequence.

ACCESSION AL136973

VERSION AL136973.7 GI:9650520

KEYWORDS HTG.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 62443)

AUTHORS Smalley,C.

TITLE Direct Submission

JOURNAL Submitted (22-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

requests: clonerequest@sanger.ac.uk

COMMENT On Aug 1, 2000 this sequence version replaced gi:9588480.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr6>
RP3-41408 is from the library RPCI-3 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/>

VECTOR: pCYPAC2

IMPORTANT: This sequence is not the entire insert of clone RP3-41408. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true right end of clone RP3-41408 is at 62443 in this sequence. The true left end of clone RP1-125E8 is at 23727 in this sequence. The true right end of clone RP4-753D5 is at 100 in this sequence.

FEATURES

	source
repeat_region	1..62443
repeat_region	/organism="Homo sapiens"
repeat_region	/db_xref="taxon:9606"
repeat_region	/chromosome="6"
repeat_region	/clone="RP3-41408"
repeat_region	/clone_lib="RPCI-3"
repeat_region	1..1203
repeat_region	/note="TIGER1 repeat: matches 1204..2417 of consensus"
repeat_region	1208..1613
repeat_region	/note="LIMC4 repeat: matches 7432..7738 of consensus"
repeat_region	1614..1916
repeat_region	/note="AluX repeat: matches 5..304 of consensus"
repeat_region	1917..2186
repeat_region	/note="LIMC4 repeat: matches 7155..7432 of consensus"
repeat_region	3084..3145
repeat_region	/note="L2 repeat: matches 2683..2744 of consensus"
repeat_region	3253..3418
repeat_region	/note="MER20 repeat: matches 40..216 of consensus"
repeat_region	3433..3493
repeat_region	/note="MIR repeat: matches 35..95 of consensus"
repeat_region	complement(3697..4051)
repeat_region	/note="match: GSS: Em:AQ018541"
repeat_region	4461..4557
repeat_region	/note="MER5A repeat: matches 11..110 of consensus"
repeat_region	5153..5364
repeat_region	/note="MER3 repeat: matches 1..209 of consensus"
repeat_region	5814..6100
repeat_region	/note="AluSq repeat: matches 1..287 of consensus"
repeat_region	6101..6403
repeat_region	/note="AluJo repeat: matches 3..304 of consensus"
repeat_region	7510..7741
repeat_region	/note="MIR repeat: matches 9..231 of consensus"
repeat_region	7962..8043
repeat_region	/note="41 copies 2 mer ta 79% conserved"
repeat_region	8047..8284
repeat_region	/note="MIR repeat: matches 2..250 of consensus"
repeat_region	8915..8973
repeat_region	/note="L2 repeat: matches 2688..2748 of consensus"
repeat_region	9781..9928
repeat_region	/note="MER5A repeat: matches 38..183 of consensus"
repeat_region	complement(11696..11949)
repeat_region	/note="match: GSS: Em:AQ079494"
repeat_region	12517..13266
repeat_region	/note="L2 repeat: matches 1947..2742 of consensus"

repeat_region	13351..13626	/note="LIMC1 repeat: matches 6055..6332 of consensus"
repeat_region	13800..13972	/note="MER41B repeat: matches 258..427 of consensus"
repeat_region	13974..14271	/note="AluX repeat: matches 2..300 of consensus"
repeat_region	14503..15494	/note="LIMC1 repeat: matches 5032..6055 of consensus"
repeat_region	15499..16381	/note="L2 repeat: matches 874..1794 of consensus"
repeat_region	18235..18540	/note="AluSP repeat: matches 4..310 of consensus"
repeat_region	18783..19092	/note="AluSg repeat: matches 1..306 of consensus"
repeat_region	21018..21135	/note="FIAM.A repeat: matches 1..125 of consensus"
repeat_region	21787..22865	/note="L2 repeat: matches 1597..2748 of consensus"
repeat_region	23139..23328	/note="LIPB1 repeat: matches 5962..6155 of consensus"
repeat_region	23397..23497	/note="MIR repeat: matches 3..111 of consensus"
repeat_region	23747..24114	/note="match: STS: Em:HS125B8T"
repeat_region	24914..25044	/note="MIR repeat: matches 20..155 of consensus"
repeat_region	25634..25712	/note="L2 repeat: matches 2631..2710 of consensus"
repeat_region	25984..26110	/note="L2 repeat: matches 2621..2750 of consensus"
repeat_region	26431..26558	/note="MIR repeat: matches 25..154 of consensus"
repeat_region	27082..27282	/note="MLTJ repeat: matches 70..271 of consensus"
repeat_region	27401..27436	/note="18 copies 2 mer tt 91% conserved"
repeat_region	27802..28277	/note="MIR repeat: matches 1..496 of consensus"
repeat_region	28298..28454	/note="MIR repeat: matches 81..262 of consensus"
repeat_region	31015..31121	/note="L2 repeat: matches 2632..2736 of consensus"
repeat_region	31149..31188	/note="20 copies 2 mer ga 95% conserved"
repeat_region	31193..31367	/note="L2 repeat: matches 2446..2618 of consensus"
repeat_region	31603..31894	/note="AluX repeat: matches 1..291 of consensus"
repeat_region	32574..32666	/note="LIMC/D repeat: matches 5726..5818 of consensus"
repeat_region	32967..33022	/note="LIMC1 repeat: matches 4943..4998 of consensus"
repeat_region	33049..33502	/note="LIMC1 repeat: matches 4477..4928 of consensus"
repeat_region	34386..34463	/note="MER91C repeat: matches 56..140 of consensus"
repeat_region	35481..35764	/note="AluSg repeat: matches 1..301 of consensus"
repeat_region	complement(36545..37035)	/note="match: GSS: Em:AQ792041"
repeat_region	37154..37502	/note="match: GSS: Em:AQ100693"
repeat_region	37699..37902	/note="HAL1 repeat: matches 707..926 of consensus"
repeat_region	38154..38291	/note="LIM4 repeat: matches 5294..5429 of consensus"
repeat_region	38292..38855	/note="LIM4 repeat: matches 5493..6068 of consensus"
repeat_region	38856..39155	/note="AluSg repeat: matches 11..303 of consensus"
repeat_region	39156..39389	/note="LIM4 repeat: matches 6068..6298 of consensus"
repeat_region	39707..39818	

```

/Note="match: GSS: Em:AQ505447"
39828..39869
/Note="21 copies 2 mer ta 76% conserved"
40072..40357
/Note="AluJo repeat: matches 1..299 of consensus"
40513..40737
/Note="MER30 repeat: matches 1..230 of consensus"
42016..42225
/Note="MIR repeat: matches 46..252 of consensus"
43759..43835
/Note="MIR repeat: matches 58..140 of consensus"
44266..44828
/Note="match: GSS: Em:AQ801525"
45767..45904
/Note="MER5A repeat: matches 27..187 of consensus"
45972..46249
/Note="MER58B repeat: matches 1..305 of consensus"
46323..46519
/Note="MIR repeat: matches 49..251 of consensus"
47083..47145
/Note="MIR repeat: matches 41..101 of consensus"
47155..47376
/Note="L1P8 repeat: matches 5925..6163 of consensus"
47773..48171
/Note="L1R16B repeat: matches 52..464 of consensus"
48504..48843
/Note="MER44A repeat: matches 8..333 of consensus"
49794..49906
/Note="MIR repeat: matches 134..261 of consensus"
49988..50103
/Note="58 copies 2 mer aa 60% conserved"
complement(50440..50936)
/Note="match: GSS: Em:AQ664348"
50614..50736

Query Match
Best Local Similarity 100.0%; Score 63; DB 9; Length 62443;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 335 GTCTCTACTTAAATACAAAAATTAGCCAGGTGGTGGGCACACGCTGTAGTCCAGCT 394
Db 14161 GTCTCTACTTAAATACAAAAATTAGCCAGGTGGTGGGCACACGCTGTAGTCCAGCT 14102

QY 395 ACT 397
|||
Db 14101 ACT 14099

RESULT 36
HSY237C10_3
WPCOMMENT
Sequence split into 4 fragments LOCUS HSY237C10 Accession AL031601
Fragment Name Begin End
HSY237C10_0 1 110000
HSY237C10_1 100001 210000
HSY237C10_2 200001 310000
HSY237C10_3 300001 385713
Continuation (4 of 4) of HSY237C10 from base 300001 (AL031601 Human DNA sequence from ol
zinc finger protein, a pseudogene similar to C.e1.

Query Match
Best Local Similarity 100.0%; Score 63; DB 9; Length 85713;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 335 GTCTCTACTTAAATACAAAAATTAGCCAGGTGGTGGGCACACGCTGTAGTCCAGCT 394
Db 52804 GTCTCTACTTAAATACAAAAATTAGCCAGGTGGTGGGCACACGCTGTAGTCCAGCT 52863

QY 395 ACT 397
|||
Db 52864 ACT 52866

```

RESULT 37
AL844148
LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AL844148 160010 bp DNA linear HTG 17-AUG-2002
Homo sapiens chromosome 1 clone RP11-104116, *** SEQUENCING IN
PROGRESS ***, 8 unordered pieces.
AL844148
AL844148.2 GI:22416202
HTG; HTGS_PHASE1; HTGS_DRAFT.
human.
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 160010)
McLay, K.

Direct Submission
Submitted (16-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Aug 21, 2002 this sequence version replaced gi:21912328.

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC

Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

----- Project Information
Center project name: BA104116

----- Summary Statistics
Assembly program: XGAP4; version 4.5

Chemistry: Dye-terminator; 100% of reads

Consensus quality: 158580 bases at least Q40

Consensus quality: 158916 bases at least Q30

Consensus quality: 159122 bases at least Q20

Insert size: 159310; sum-of-contigs

Insert size: 173409; 4.8% error; agarose-fp

Quality coverage: 6.64x in Q20 bases; sum-of-contigs Quality
coverage: 6.22x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 3853: contig of 3853 bp in length
3854 3953: gap of 100 bp
3954 24930: contig of 20977 bp in length
24931 25030: gap of 100 bp
25031 32722: contig of 7692 bp in length
32723 32822: gap of 100 bp
32823 72051: contig of 39229 bp in length
72052 72151: gap of 100 bp
72152 81625: contig of 9474 bp in length
81626 81725: gap of 100 bp
81726 103210: contig of 21485 bp in length
103211 103310: gap of 100 bp
103311 127637: contig of 24327 bp in length
127638 127737: gap of 100 bp
127738 160010: contig of 32273 bp in length.

Location/Qualifiers

1..160010

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="1"

/clone="RP11-104116"

/clone_lib="RPC1-11.1"

1..3853

/note="assembly_fragment:01821"

fragment_chain:1

3954..24930

/note="assembly_fragment:01665"

fragment_chain:1

FEATURES

source

misc_feature

misc_feature

misc_feature 25031..32722 /note="assembly_fragment:00578
fragment_chain:1"
32823..72051 /note="assembly_fragment:00760
fragment_chain:1"
72152..81625 /note="assembly_fragment:00250
fragment_chain:1"
81726..103210 /note="assembly_fragment:01744
fragment_chain:1"
103311..127637 /note="assembly_fragment:01572
fragment_chain:1"
127738..160010 /note="assembly_fragment:00210
fragment_chain:1"
clone_end:T7
vector_side:right
BASE COUNT 48388 a 32237 c 32257 g 46428 t 700 others
ORIGIN

Query Match 6.3%; Score 63; DB 2; Length 160010;
Best Local Similarity 100.0%; Pred. No. 2.3e-25;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 335 GTCTCTACTATAAATACAAAATAGCCAGGTGGTGGCACACGCGTGTAGTCCAGCT 394
|||||
DB 97633 GTCTCTACTATAAATACAAAATAGCCAGGTGGTGGCACACGCGTGTAGTCCAGCT 97692
QY 395 ACT 397
|||
DB 97693 ACT 97695

RESULT 38
AL391479
LOCUS
DEFINITION Homo sapiens chromosome 10 clone RP13-268C11, *** SEQUENCING IN
PROGRESS ***, 8 unordered pieces.
ACCESSION AL391479
VERSION AL391479.3 GI:10120141
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Sims.S.
Direct Submission
Submitted (20-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Sep 13, 2000 this sequence version replaced gi:9864729.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BB268C11
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 164822 bases at least Q40
Consensus quality: 166195 bases at least Q30
Consensus quality: 166903 bases at least Q20
Insert size: 167371; sum-of-contigs
Insert size: 168395; 7.1% error; agarose-fp
Quality coverage: 4.97x in Q20 bases; sum-of-contigs Quality
coverage: 5.02x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 6683: contig of 6683 bp in length
* 6684 6783: gap of 100 bp
* 6784 89821: contig of 83038 bp in length
* 89822 89921: gap of 100 bp
* 89922 92032: contig of 2111 bp in length
* 92033 92132: gap of 100 bp
* 92133 99119: contig of 6987 bp in length
* 99120 99219: gap of 100 bp
* 99220 105974: contig of 6755 bp in length
* 105975 106074: gap of 100 bp
* 106075 144349: contig of 38275 bp in length
* 144350 144449: gap of 100 bp
* 144450 150629: contig of 6180 bp in length
* 150630 150729: gap of 100 bp
* 150730 168071: contig of 17342 bp in length.
FEATURES
Location/Qualifiers
1..168071
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP13-268C11"
/clone_lib="RPC1-13.2"
1..6683
/note="assembly_fragment:00272
fragment_chain:1
clone_end:SP6
vector_side:left"
6784..89821
/note="assembly_fragment:01884
fragment_chain:1"
89922..92032
/note="assembly_fragment:01743
fragment_chain:1"
92133..99119
/note="assembly_fragment:00870
fragment_chain:1"
99220..105974
/note="assembly_fragment:02095
fragment_chain:1"
106075..144349
/note="assembly_fragment:00224
fragment_chain:2"
144450..150629
/note="assembly_fragment:01619
fragment_chain:2"
150730..168071
/note="assembly_fragment:01282
fragment_chain:2"
BASE COUNT 50488 a 34118 c 34260 g 48505 t 700 others
ORIGIN

Query Match 6.3%; Score 63; DB 2; Length 168071;
Best Local Similarity 100.0%; Pred. No. 2.3e-25;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 335 GTCTCTACTATAAATACAAAATAGCCAGGTGGTGGCACACGCGTGTAGTCCAGCT 394
|||||
DB 43567 GTCTCTACTATAAATACAAAATAGCCAGGTGGTGGCACACGCGTGTAGTCCAGCT 43626
QY 395 ACT 397
|||
DB 43627 ACT 43629

RESULT 39

AC021761/c
 LOCUS
 DEFINITION Homo sapiens chromosome 11, clone RP11-361M6, linear PRI 22-MAR-2002
 AC021761
 ACCESSION
 VERSION
 AC021761.8 GI:19570169
 HTG.
 SOURCE
 Homo sapiens.
 ORGANISM

REFERENCE

1 (bases 1 to 182789)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Unpublished

TITLE

Homo sapiens chromosome 11, clone RP11-361M6

REFERENCE

2 (bases 1 to 182789)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F., Boguslavskiy,L., Boukhalter,B., Brown,A., Burkett,G., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArelano,K., Dewar,K., Domino,M., Doyle,M., Feneator,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Headford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
 Direct Submission

TITLE

Submitted (20-JAN-2000) Whitehead Institute/MIT Center for Genome

JOURNAL

Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 182789)

REFERENCE

Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Lakocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., MacDonald,P., Major,J., Marquis,N., Matthews,C., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phukhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission

TITLE

Submitted (21-MAR-2002) Whitehead Institute/MIT Center for Genome

JOURNAL

Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 182789)

REFERENCE

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Lakocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,

MacLean,C., MacDonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phukhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission

TITLE

Submitted (22-MAR-2002) Whitehead Institute/MIT Center for Genome

JOURNAL

Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Mar 21, 2002 this sequence version replaced gi:15390958.
 All repeats were identified using RepeatMasker:

COMMENT

Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RN/RepeatMasker.html

REFERENCE

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence.submissions@genome.wi.mit.edu

PROJECT

----- Project Information

CENTER

Center project name: L2173

CLONE

Center clone name: 361_M_6

FEATURES

Location/Qualifiers

1. .182789
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11"
 /clone="RP11-361M6"
 /clone.lib="RPC1-11 Human Male BAC"
 complement(23..432)
 /rpt_family="L1Mcb"
 complement(436..751)
 /rpt_family="L1Mcb"
 complement(777..1261)
 /rpt_family="L1M4"
 complement(1291..2396)
 /rpt_family="L1M4"
 complement(2397..2773)
 /rpt_family="MSTA"
 complement(2774..4396)
 /rpt_family="MSTA-int"
 complement(4399..4760)
 /rpt_family="MSTA"
 complement(4797..4930)
 /rpt_family="L1M4"
 4931..4953
 /rpt_family="(CARAC)n"
 complement(4954..5413)
 /rpt_family="L1M4"
 complement(5610..6614)
 /rpt_family="L1Mcb"
 complement(6628..7091)
 /rpt_family="L1Mca"
 complement(7092..7826)
 /rpt_family="LTR8"
 complement(7827..8358)
 /rpt_family="L1Mca"
 8428..8583
 /rpt_family="L1MEL"
 complement(8584..8890)
 /rpt_family="AluSq"
 8891..9083
 /rpt_family="L1MEL"
 9179..9293
 /rpt_family="L1MEL"
 10399..10454

```

/rpt_family="AT-rich"
complement(10680..10828)
/rpt_family="MIR"
11360..11388
/rpt_family="AT-rich"
complement(11390..12899)
/rpt_family="LIP3A"
13179..13541
/rpt_family="LIME1"
13553..13654
/rpt_family="LIMA10"
13799..14122
/rpt_family="MLT1A1"
14955..14976
/rpt_family="AT-rich"
15127..15266
/rpt_family="MIR"
15402..15578
/rpt_family="L2"
16410..16722
/rpt_family="MER58B"
17098..17292
/rpt_family="MLT1A1"
complement(18723..18931)
/rpt_family="MLT1A2"
complement(19010..19352)
/rpt_family="THE1B"
complement(19354..19392)
/rpt_family="THE1B-int"
complement(19376..20611)
/rpt_family="THE1B-int"
complement(20612..20966)
/rpt_family="THE1B"
complement(20967..21061)
/rpt_family="MLT1A2"
21122..21316
/rpt_family="MER58A"
complement(21376..21572)
/rpt_family="MER5A"
complement(21883..22035)
/rpt_family="MIR"
22330..22428
/rpt_family="L3"
complement(22475..22680)
/rpt_family="MIR"
23473..23513
/rpt_family="(TG)n"
complement(24066..24302)
/rpt_family="MIR"
complement(24586..24747)
/rpt_family="Charlie8"
25296..25316
/rpt_family="(TA)n"
27357..27378
/rpt_family="AT-rich"
27813..27944

Query Match      6.3%; Score 63; DB 9; Length 182789;
Best Local Similarity 100.0%; Pred.No. 2.3e-25;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 343 TAAATACAAAATAGCCAGGTGTGGTGGCAGCAGCTGTAGTCCAGCTACTTGGGA 402
|||||
Db 65095 TAAATACAAAATAGCCAGGTGTGGTGGCAGCAGCTGTAGTCCAGCTACTTGGGA 65036
|||||

QY 403 GGC 405
|||
Db 65035 GGC 65033

RESULT 40
AL354975/c
LOCUS          AL354975          212730 bp    DNA    linear    HTG 11-SEP-2001

```

```

DEFINITION Homo sapiens chromosome 10 clone RP11-300L24, *** SEQUENCING IN
PROGRESS ***, 8 unordered pieces.
ACCESSION AL354975
VERSION AL354975.15 GI:15590902
KEYWORDS HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 212730)
AUTHORS Phillimore,B.
TITLE Direct Submission
JOURNAL Submitted (10-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
COMMENT requests: clonerequest@sanger.ac.uk
On Sep 12, 2001 this sequence version replaced gi:14586088.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA300L24
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 211054 bases at least Q40
Consensus quality: 211410 bases at least Q30
Consensus quality: 211660 bases at least Q20
Insert size: 212030; sum-of-contigs
Insert size: 197391; 9.1% error; agarose-fp
Quality coverage: 8.73x in Q20 bases; sum-of-contigs Quality
coverage: 9.68x in Q20 bases; agarose-fp
-----
* NOTE: this is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 9151: contig of 9151 bp in length
* 9152 9251: gap of 100 bp
* 9252 41384: contig of 32133 bp in length
* 41385 41484: gap of 100 bp
* 41485 54463: contig of 12979 bp in length
* 54464 54563: gap of 100 bp
* 54564 56880: contig of 2317 bp in length
* 56881 56980: gap of 100 bp
* 56981 77886: contig of 20905 bp in length
* 77887 122287: contig of 44301 bp in length
* 122288 122387: gap of 100 bp
* 122388 128570: contig of 6183 bp in length
* 128571 128670: gap of 100 bp
* 128671 212730: contig of 84060 bp in length.
*
FEATURES
Location/Qualifiers
1..212730
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-300L24"
/clone.lib="RPCT-11.2"
1..9151
/misc_feature
/note="assembly_fragment:03340
fragment_chain:1
clone_end:T7
vector_side:left"
9252..41384
/misc_feature
/note="assembly_fragment:02020
fragment_chain:1"

```

```

misc_feature 41485..54463
/note="assembly_fragment:00892
fragment_chain:2"
misc_feature 54564..56880
/note="assembly_fragment:04482
fragment_chain:2"
misc_feature 56981..77886
/note="assembly_fragment:02718
fragment_chain:2"
misc_feature 77987..122287
/note="assembly_fragment:03863
fragment_chain:2"
misc_feature 122388..128570
/note="assembly_fragment:02503
fragment_chain:2"
misc_feature 128671..212730
/note="assembly_fragment:01108
fragment_chain:2"
clone_end:sp6
vector_side:right"
BASE COUNT 61260 a 42923 c 43637 g 64207 t 703 others
ORIGIN
Query Match
Best Local Similarity 100.0%; Pred. No. 2.4e-25;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 335 GTCCTACTAAAATACAAAATAGCCAGGTGTGTGGCAGCGCTGTAGTCCAGCT 394
|||||
Db 211900 GTCCTACTAAAATACAAAATAGCCAGGTGTGTGGCAGCGCTGTAGTCCAGCT 211841
QY 395 ACT 397
||||
Db 211840 ACT 211838

RESULT 41
AL591130
LOCUS
DEFINITION Homo sapiens chromosome 10 clone RP13-251015, *** SEQUENCING IN
PROGRESS ***, 32 unordered pieces.
ACCESSION AL591130
VERSION AL591130.1 GI:13990326
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Burton, J.
1
Direct Submission
Submitted (18-DEC-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: B8251015
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 217028 bases at least Q40
Consensus quality: 223277 bases at least Q30
Consensus quality: 226745 bases at least Q20
Insert size: 229062; sum-of-contigs
Quality coverage: 189083; 10.2% error; agarose-fp
Quality coverage: 3.87x in Q20 bases; sum-of-contigs Quality
coverage: 5.96x in Q20 bases; agarose-fp
-----

```

* NOTE: This is a 'working draft' sequence. It currently consists of 32 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 5872: contig of 5872 bp in length
5873 5972: gap of 100 bp
5973 11302: contig of 5330 bp in length
11303 11402: gap of 100 bp
11403 15003: contig of 3601 bp in length
15004 15103: gap of 100 bp
15104 19081: contig of 3978 bp in length
19082 28512: contig of 9331 bp in length
28513 28612: gap of 100 bp
28613 36009: contig of 7397 bp in length
36010 36109: gap of 100 bp
36110 43831: contig of 7722 bp in length
43832 43931: gap of 100 bp
43932 101489: contig of 57558 bp in length
101490 101589: gap of 100 bp
101590 137416: contig of 35827 bp in length
137417 137516: gap of 100 bp
137517 150711: contig of 13195 bp in length
150712 150811: gap of 100 bp
150812 159076: contig of 8265 bp in length
159077 159176: gap of 100 bp
159177 166731: contig of 7555 bp in length
166732 166831: gap of 100 bp
166832 170446: contig of 3615 bp in length
170447 170546: gap of 100 bp
170547 183663: contig of 13117 bp in length
183664 183763: gap of 100 bp
183764 186875: contig of 3112 bp in length
186876 186975: gap of 100 bp
186976 189138: contig of 2163 bp in length
189139 189238: gap of 100 bp
189239 191366: contig of 2128 bp in length
191367 191466: gap of 100 bp
191467 193973: contig of 2507 bp in length
193974 194073: gap of 100 bp
194074 196413: contig of 2340 bp in length
196414 196513: gap of 100 bp
196514 199969: contig of 3456 bp in length
199970 200069: gap of 100 bp
200070 203326: contig of 3257 bp in length
203327 203426: gap of 100 bp
203427 206251: contig of 2825 bp in length
206252 206351: gap of 100 bp
206352 208716: contig of 2365 bp in length
208717 208816: gap of 100 bp
208817 211743: contig of 2927 bp in length
211744 211843: gap of 100 bp
211844 214458: contig of 2615 bp in length
214459 214558: gap of 100 bp
214559 216763: contig of 2205 bp in length
216764 216863: gap of 100 bp
216864 220060: contig of 3197 bp in length
220061 220160: gap of 100 bp
220161 223061: contig of 2901 bp in length
223062 223161: gap of 100 bp
223162 225305: contig of 2144 bp in length
225306 225405: gap of 100 bp
225406 227527: contig of 2122 bp in length
227528 227627: gap of 100 bp
227628 229963: contig of 2336 bp in length
229964 230063: gap of 100 bp
230064 232162: contig of 2099 bp in length.

Location/Qualifiers
1. .232162

FEATURES
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/clone_lib="RPC1-13.2"
1. 5872
/note="assembly_fragment:00817
fragment_chain:1"
5973. 11302
/note="assembly_fragment:02075
fragment_chain:1"
11403. 15003
/note="assembly_fragment:02256
fragment_chain:1"
15104. 19081
/note="assembly_fragment:00201
fragment_chain:1"
19182. 28512
/note="assembly_fragment:01282
fragment_chain:1"
28613. 36009
/note="assembly_fragment:00649
fragment_chain:2"
36110. 43831
/note="assembly_fragment:00481
fragment_chain:2"
43932. 101489
/note="assembly_fragment:01511
fragment_chain:2"
101590. 137416
/note="assembly_fragment:01201
fragment_chain:2"
137517. 150711
/note="assembly_fragment:00117
fragment_chain:3"
150812. 159076
/note="assembly_fragment:00118
fragment_chain:3"
159177. 166731
/note="assembly_fragment:01939
fragment_chain:3"
166832. 170446
/note="assembly_fragment:02774
fragment_chain:4"
170547. 183663
/note="assembly_fragment:00485
fragment_chain:4"
183764. 186875
/note="assembly_fragment:00790
fragment_chain:4"
186976. 189138
/note="assembly_fragment:01385
fragment_chain:5"
189239. 191366
/note="assembly_fragment:01448
fragment_chain:5"
191467. 193973
/note="assembly_fragment:01471
fragment_chain:6"
194074. 196413
/note="assembly_fragment:02190
fragment_chain:6"
196514. 199969
/note="assembly_fragment:00042"
200070. 203326
/note="assembly_fragment:00679"
203427. 206251
/note="assembly_fragment:01099"
206352. 208716
/note="assembly_fragment:01233"
208817. 211743
/note="assembly_fragment:01246"
211844. 214458

```

```

/misc_feature /note="assembly_fragment:01479"
214559. 216763
/misc_feature /note="assembly_fragment:01572"
216864. 220060
/misc_feature /note="assembly_fragment:01668"
220161. 223061
/misc_feature /note="assembly_fragment:01866"
223162. 225305
/misc_feature /note="assembly_fragment:02021"
225406. 227527
/misc_feature /note="assembly_fragment:02054"
227628. 229963
/misc_feature /note="assembly_fragment:02419"
230064. 232162
/misc_feature /note="assembly_fragment:02818"
230064. 232162
BASE COUNT 69322 a 46531 c 42921 g 70283 t 3105 others
ORIGIN
Query Match 6.3%; Score 63; DB 2; Length 232162;
Best Local Similarity 100.0%; Pred. No. 2.4e-25;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 335 GTCTTACTAAAAATACAAAATTAGCAGGTGTGGTGGCACACGCTGTAGTCCAGCT 394
|||||
Db 114332 GTCTTACTAAAAATACAAAATTAGCAGGTGTGGTGGCACACGCTGTAGTCCAGCT 114391
QY 395 ACT 397
|||
Db 114392 ACT 114394

RESULT 42
AP000098/c
LOCUS
DEFINITION Homo sapiens genomic DNA, chromosome 21q21.1-q21.2, clone:B304D2,
LL56-APP region, complete sequence.
ACCESSION AP000098
VERSION AP000098.2 GI:7077210
KEYWORDS HTG.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens 68,122 genomic DNA of 21q21.1-q21.2
JOURNAL Published Only in Database (1999)
REFERENCE 2 (bases 1 to 68122)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
TITLE Submitted (24-DEC-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
Kitasato Univ., 1-15-1 Kitasato, Sagami-hara, Kanagawa 228-8555,
Japan (E-mail:hattori@gsc.riken.go.jp,
URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9924)
COMMENT On Feb 24, 2000 this sequence version replaced gi:6997413.
FEATURES
Location/Qualifiers
Source
1. 68122
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q21.1-q21.2"
/clone="B304D2"
BASE COUNT 20417 a 12290 c 13407 g 22008 t
ORIGIN
Query Match 6.2%; Score 62; DB 9; Length 68122;
Best Local Similarity 100.0%; Pred. No. 8.6e-25;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 336 TCTCTACTAAATAACAAAAATTAGCCAGGTGTGGTGGCACACGCCCTGTAGTCCAGCTA 395
|||||
Db 21283 TCTCTACTAAATAACAAAAATTAGCCAGGTGTGGTGGCACACGCCCTGTAGTCCAGCTA 21224
|||||
QY 396 CT 397
||
Db 21223 CT 21222

RESULT 43
AC025810 71530 bp DNA linear PRI 04-JAN-2002
LOCUS Homo sapiens chromosome 16 clone CTD-2034121, complete sequence.
DEFINITION AC025810
ACCESSION AC025810
VERSION AC025810.4 GI:18057077
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 71530)
DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 71530)
DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 71530)
DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (04-JAN-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jan 4, 2002 this sequence version replaced gi:14589449.
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.
-----
FEATURES
source Location/Qualifiers
1..71530
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/clone="CTD-2034121"
BASE COUNT 21125 a 16392 c 15391 g 18622 t
ORIGIN
Query Match 6.2%; Score 62; DB 9; Length 71530;
Best Local Similarity 100.0%; Pred. No. 8.6e-25;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 336 TCTCTACTAAATAACAAAAATTAGCCAGGTGTGGTGGCACACGCCCTGTAGTCCAGCTA 395
|||||
Db 65905 TCTCTACTAAATAACAAAAATTAGCCAGGTGTGGTGGCACACGCCCTGTAGTCCAGCTA 65964
|||||
QY 396 CT 397

RESULT 45
AC079855 143146 bp DNA linear PRI 01-MAR-2002
LOCUS Homo sapiens BAC clone RP11-332116 from 7, complete sequence.
DEFINITION AC079855
ACCESSION AC079855
VERSION AC079855.8 GI:14717359
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 143146)
Sulston, J.E. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)

```

```

Db 65965 CT 65966
||
RESULT 44
AC016593 124211 bp DNA linear PRI 07-NOV-2000
LOCUS Homo sapiens chromosome 5 clone CTD-2005H4, complete sequence.
DEFINITION AC016593
ACCESSION AC016593
VERSION AC016593.5 GI:111119456
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 124211)
DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 124211)
DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (04-DEC-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 124211)
DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (07-NOV-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Nov 7, 2000 this sequence version replaced gi:7710203.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.sngc.stanford.edu
Quality: Phrap Quality >=40 99.5% of Sequence;
Estimated Total Number of Errors is 0.6.
FEATURES
source Location/Qualifiers
1..124211
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2005H4"
BASE COUNT 37357 a 22562 c 23975 g 40317 t
ORIGIN
Query Match 6.2%; Score 62; DB 9; Length 124211;
Best Local Similarity 100.0%; Pred. No. 8.9e-25;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 336 TCTCTACTAAATAACAAAAATTAGCCAGGTGTGGTGGCACACGCCCTGTAGTCCAGCTA 395
|||||
Db 43273 TCTCTACTAAATAACAAAAATTAGCCAGGTGTGGTGGCACACGCCCTGTAGTCCAGCTA 43214
|||||
QY 396 CT 397
||
Db 43213 CT 43212

RESULT 45
AC079855/c 143146 bp DNA linear PRI 01-MAR-2002
LOCUS Homo sapiens BAC clone RP11-332116 from 7, complete sequence.
DEFINITION AC079855
ACCESSION AC079855
VERSION AC079855.8 GI:14717359
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 143146)
Sulston, J.E. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)

```

99063792
 9847074
 2 (bases 1 to 143146)
 Mclella, M. and Kozlowski, A.
 The sequence of Homo sapiens BAC clone RP11-332L16
 Unpublished (2001)
 3 (bases 1 to 143146)
 Waterston, R.H.
 Direct Submission
 Submitted (14-SEP-2000) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 4 (bases 1 to 143146)
 Waterston, R.H.
 Direct Submission
 Submitted (13-JUL-2001) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 5 (bases 1 to 143146)
 Waterston, R.
 Direct Submission
 Submitted (01-MAR-2002) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Jul 13, 2001 this sequence version replaced gi:13794255.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@watson.wustl.edu
 ----- Summary Statistics

 Center project name: H_NH0332L16

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and
 sequencing collaboration between the NHGRI Chromosome 7 Mapping
 Project (Eric D. Green, Director), John D. McPherson in the
 Department of Genetics (Washington University), and the Washington
 University Genome Sequencing Center. For additional information
 about the map position of this sequence, see
<http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send
 mailto:egreen@nhgri.nih.gov, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male
 donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frenken, E.,
 Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved
 approach for construction of bacterial artificial chromosome
 libraries. Genomics 51:1-8. The clone may be obtained either from
 Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong
 and coworkers at <http://www.chori.org>
 VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-379F11; the clone sequenced
 to the right is RP5-1132H15, 2000 bp overlap. Actual start of this
 clone is at base position 1 of RP11-332L16; actual end is at base
 position 31561 of RP5-1132H15.

Data from AC024641 and AC026281 was used to finish this clone,

AC079855. Polymorphisms exist between RP11-332L16 and RP5-1132H15.
 The sequence between 46100-46104 is a single chemistry region.
 Location/Qualifiers
 1. 143146
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7"
 /clone="RP11-332L16"
 /clone_lib="RPC1-11"
 81. 434
 /rpt_family="L1"
 315. 336
 /rpt_family="AT_rich"
 437. 494
 /rpt_family="MER1_type"
 496. 793
 /rpt_family="L2"
 838. 1377
 /rpt_family="L1"
 1428. 1661
 /rpt_family="L1"
 2283. 2481
 /rpt_family="L2"
 2939. 2959
 /rpt_family="(A)n"
 3711. 3976
 /rpt_family="Alu"
 3815. 3840
 /rpt_family="AT_rich"
 4114. 4422
 /rpt_family="Alu"
 4441. 4601
 /rpt_family="L1"
 4720. 5015
 /rpt_family="L1"
 5001. 5040
 /rpt_family="AT_rich"
 5466. 5631
 /rpt_family="L1"
 6107. 6396
 /rpt_family="Alu"
 6413. 6710
 /rpt_family="Alu"
 12767. 12817
 /rpt_family="(TAAAA)n"
 12922. 13125
 /note="match to EST AL562533 (NID:gl2911047)"
 12925. 13125
 /note="match to EST AAL28112 (NID:gl687996) z123g08.rl"
 12925. 13125
 /note="match to EST AA445921 (NID:g2158586) zW57f04.rl"
 12925. 13125
 /note="match to EST AL550344 (NID:gl2887227)"
 12925. 13125
 /note="similar to Mus musculus EST W64675 (NID:gl372592)
 me08d02.rl"
 12925. 13097
 /note="match to EST AL523147 (NID:gl2786640)"
 12984. 13125
 /note="match to EST BG121389 (NID:gl2614898)"
 13137. 13170
 /rpt_family="AT_rich"
 13530. 13630
 /rpt_family="L1"
 14292. 14295
 /note="match to EST AAL28112 (NID:gl687996) z123g08.rl"
 14325. 14422
 /rpt_family="Achobo"
 14428. 14519
 /rpt_family="MER2_type"
 14819. 15084
 /rpt_family="Alu"

```

repeat_region 15063..15084
                /rpt_family="(A)n"
repeat_region 15675..15801
                /rpt_family="L2"
repeat_region 15912..16242
                /rpt_family="MERL_type"
repeat_region 16353..16648
                /rpt_family="Alu"
repeat_region 16827..16887
                /rpt_family="L2"
repeat_region 17149..17356
                /rpt_family="MERL_type"
repeat_region 17956..18100
                /rpt_family="L1"
repeat_region 18114..18688
                /rpt_family="L1"
repeat_region 18927..19255
                /rpt_family="L1"
misc_feature 19300..19539
                /note="similar to Homo sapiens EST B6718062
                (NID:gi3997249)"
misc_feature 19573..19847
                /note="similar to Homo sapiens EST BF512467
                (NID:gi1597569)"
misc_feature 19687..19814
                /note="similar to Homo sapiens EST B6718062
                (NID:gi3997249)"
repeat_region 20071..20704
                /rpt_family="L1"
repeat_region 20826..21132
                /rpt_family="Alu"
repeat_region 21191..21506
                /rpt_family="L2"
repeat_region 21507..21535
                /rpt_family="(T)n"
repeat_region 21771..21929
                /rpt_family="L2"
repeat_region 23083..23251
                /rpt_family="Alu"
repeat_region 23245..23290
                /rpt_family="AT-rich"

```

```

Query Match      6.2%; Score 62; DB 9; Length 143146;
Best Local Similarity 100.0%; Pred. No. 9e-25;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 245 CCTGTAATCCAGCAGCTTGGGAGGCTGAGACGGGTGGATCACCTGAAGTCAGGAGTTCA 304
    |||||
Db 136003 CCTGTAATCCAGCAGCTTGGGAGGCTGAGACGGGTGGATCACCTGAAGTCAGGAGTTCA 135944

Qy 305 AG 306
    ||
Db 135943 AG 135942

```

Search completed: November 24, 2002, 12:31:27
 Job time : 7279 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2002, 13:27:24 ; Search time 1549.5 seconds
(without alignments)
261.302 Million cell updates/sec

Title: US-09-784-423-124

Perfect score: 25

Sequence: 1 GGTTCAGTGCACCGAGATAAGACT 25

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*

1: em_estba.*
2: em_esthu.*
3: em_estin.*
4: em_estnu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_esti.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_man.*
24: em_gss_mus.*
25: em_gss_rus.*
26: em_gss_oth.*
27: em_gss_rtd.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	25	100.0	337	17	AQ356083
C 2	21	84.0	166	12	BF757007
C 3	21	84.0	166	17	B85546
C 4	21	84.0	220	9	AA297398
C 5	21	84.0	281	9	AU043112
C 6	21	84.0	325	9	AA593502

C 7	21	84.0	389	14	T92723
C 8	21	84.0	397	12	BF754298
C 9	21	84.0	452	17	AQ384259
C 10	21	84.0	461	17	AQ702820
C 11	21	84.0	476	10	AW751650
C 12	21	84.0	481	10	AW265138
C 13	21	84.0	503	10	BE646402
C 14	21	84.0	509	17	AQ721912
C 15	21	84.0	514	17	B37536
C 16	21	84.0	529	17	AQ819983
C 17	21	84.0	539	17	AQ608587
C 18	21	84.0	542	13	BI494061
C 19	21	84.0	570	17	AQ708886
C 20	21	84.0	573	13	BI494060
C 21	21	84.0	575	17	AQ520887
C 22	21	84.0	649	17	AQ482581
C 23	21	84.0	666	17	AQ039496
C 24	21	84.0	668	17	AG050903
C 25	21	84.0	882	12	BG214939
C 26	21	84.0	914	12	BF032499
C 27	20	80.0	132	12	BF837405
C 28	20	80.0	139	12	BF878448
C 29	20	80.0	144	9	AA709114
C 30	20	80.0	154	9	AI626093
C 31	20	80.0	173	9	AA046824
C 32	20	80.0	175	17	B73330
C 33	20	80.0	184	9	AA348526
C 34	20	80.0	185	14	BQ251032
C 35	20	80.0	189	14	H02828
C 36	20	80.0	191	14	R44194
C 37	20	80.0	194	9	AI241886
C 38	20	80.0	201	14	H03719
C 39	20	80.0	211	14	N73648
C 40	20	80.0	212	10	AW772536
C 41	20	80.0	231	9	AA228262
C 42	20	80.0	233	10	B261473
C 43	20	80.0	234	12	BF962750
C 44	20	80.0	237	9	AA380519
C 45	20	80.0	239	17	B87492

ALIGNMENTS

RESULT 1
AQ356083/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AQ356083
CITBI-El-2534J20.TF CITBI-El Homo sapiens genomic clone 2534J20,
DNA sequence.
AQ356083
GSS.
human.
Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 337)
Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
Map Building
Unpublished (1997)

Other GSSs: CITBI-El-2534J20.TR
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.


```

Seq primer: M13-21
Class: BAC ends.
FEATURES
  source
    Location/Qualifiers
      1..337
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="2534320"
        /clone_lib="CITBI-E1"
        /sex="male"
        /cell_type="sperm"
        /note="Vector: pBelOAc11; Site_1: EcoRI; Site_2: EcoRI;
          Caltech Human BAC Library D"
BASE COUNT      78 a      70 c      75 g      114 t
ORIGIN
      1 GGTTCAGTGCAGCCGAGATAAGT 25
      249 GGTTCAGTGCAGCCGAGATAAGT 225

Query Match      100.0%; Score 25; DB 17; Length 337;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGATAAGT 25
      |||
Db 249 GGTTCAGTGCAGCCGAGATAAGT 225

RESULT 2
LOCUS      BF757007/c
DEFINITION      MR0-CT0451-021100-017-g07 CT0451 Homo sapiens cDNA, mRNA sequence.
ACCESSION      BF757007
VERSION      BF757007.1 GI:12104907
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the PAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR0&t2=MR0-CT0451-
021100-017-g07&t3=2000-11-02&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 91.
FEATURES
  source
    Location/Qualifiers
      1..166
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone_lib="CT0451"
        /dev_stage="Adult"
        /note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
          SnaI; A mini-library was made by cloning products derived
          from ORESTES PCR (U.S. Letters Patent application No. 196
          716 - Ludwig Institute for Cancer Research) profiles
          into the pUC 18 vector. Reverse transcription of tissue
          mRNA and cDNA amplification were performed under low
          stringency conditions."

```

```

BASE COUNT      39 a      32 c      35 g      60 t
ORIGIN
  Query Match      84.0%; Score 21; DB 12; Length 166;
  Best Local Similarity 100.0%; Pred. No. 0.039;
  Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGATAA 21
      |||
Db 145 GGTTCAGTGCAGCCGAGATAA 125

RESULT 3
LOCUS      B85546/c
DEFINITION      RPCI11-19J11.TP RPCI-11 Homo sapiens genomic clone RPCI-11-19J11,
DNA sequence.
ACCESSION      B85546
VERSION      B85546.1 GI:2926678
KEYWORDS      GSS.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 166)
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
Venter,J.C.
Use of BAC End Sequences for Sequence-Ready Map Building (1998)
Unpublished (1998)
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mda@imgr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@imgr.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.
FEATURES
  source
    Location/Qualifiers
      1..166
        /organism="Homo sapiens"
        /db_xref="GDB:7507138"
        /db_xref="taxon:9606"
        /clone="RPCI-11-19J11"
        /clone_lib="RPCI-11"
        /sex="Male"
        /cell_type="Lymphocytes"
        /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
          RPCI11 Human Male BAC Library"
BASE COUNT      29 a      44 c      45 g      48 t
ORIGIN
  Query Match      84.0%; Score 21; DB 17; Length 166;
  Best Local Similarity 100.0%; Pred. No. 0.039;
  Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGATAA 21
      |||
Db 165 GGTTCAGTGCAGCCGAGATAA 145

RESULT 4
LOCUS      AA297398/c
DEFINITION      EST112932 Epithelioid sarcoma Homo sapiens cDNA 5' end similar to
AA297398
EST112932 EST containing Alu repeat, mRNA sequence.

```

ACCESSION AA297398
 VERSION AA297398.1 GI:1949731
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE

AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
 , C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
 , O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,
 Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
 , L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A.,
 Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M.,
 Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
 Moreno-Palancas,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
 Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
 Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
 Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hudson,P., Kim,A.K., Kozak,D.L.,
 Hu,J.S., Greene,J.M., Gruber,J., Gruber,J., Kim,A.K., Kozak,D.L.,
 Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L.,
 Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon
 , M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
 Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence

JOURNAL

Nature 377 (6547 Suppl), 3-174 (1995)

MEDLINE

96026280

COMMENT

Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: arkerlavet@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seq primer: M13 Reverse.

FEATURES

source

1. .220

/organism="Homo sapiens"
 /db_xref="AYCC (inhost):137964"
 /db_xref="taxon:9606"
 /clone_lib="Epithelioid sarcoma"
 /dev_stage="adult"
 /note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI
 ; Site_2: XhoI"

BASE COUNT 39 a 64 c 51 g 62 t 4 others

Query Match

Best Local Similarity 84.0%; Score 21; DB 9; Length 220;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGATAA 21

|||||

Db 85 GGTTCAGTGCAGCCGAGATAA 65

RESULT 5

AU043112/c

LOCUS

DEFINITION

AU043112 Mouse sixteen-cell-embryo cDNA Mus musculus cDNA clone

J0900A09 3', mRNA sequence.

ACCESSION AU043112

VERSION AU043112.1 GI:3978587

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 281)

Ko,M.S.H., Kitchen,J.R., Wang,X., Threat,T.A., Sun,T., Depalma,G.E.,
 Liang,Y., Kargul,G.J., Sharara,R., Lim,M.K. and Doi,H.
 Systematic analyses of genes expressed in 16-cell mouse embryo (The
 ERA/Doi Project at Wayne State University)
 Unpublished (1998)
 Contact: Hirofumi Doi

Doi Biosymmetry Project, ERATO

Japan Science and Technology Corporation (JST)

WBG Marine East 12F, 2-6 Nakase, Mihama-ku, Chiba 261-71, Japan

Email: hdo@doi.jst.go.jp.

Location/Qualifiers

1. .281

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="J0900A09"

/clone_lib="Mouse sixteen-cell-embryo cDNA"

/dev_stage="sixteen-cell-embryo"

/note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia

) with a modified polylinker; Site_1: Not I; Site_2: Eco

RI; 1st strand cDNA was primed with a Not I - oligo(dT)

primer [5'

TGTACCAATCTGAAGTGGAGCGCGCGAGATGGTGTGTGTGTGTGTGTGT

T 3']; double-stranded cDNA was ligated to Eco RI

adaptors (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of the modified pT7T3 vector.

RNA provided by Dr. Minoru Ko, Wayne State Univ. Library

constructed and normalized by Bento Soares and M.Fatima

Bonaldo."

BASE COUNT 51 a 87 c 69 g 73 t 1 others

Query Match 84.0%; Score 21; DB 9; Length 281;

Best Local Similarity 100.0%; Pred. No. 0.044;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGATAA 21

|||||

Db 65 GGTTCAGTGCAGCCGAGATAA 45

RESULT 6

AA593502/c

LOCUS

DEFINITION

AA593502

nm27b11.s1 NCI-CGAP Gas1 Homo sapiens cDNA clone IMAGE:1085085 3'

similar to contains Alu repetitive element; contains L1.t1 L1

repetitive element ; mRNA sequence.

AA593502

AA593502.1 GI:2409264

EST.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 325)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbbs@mail.nih.gov

Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: Stratagene, Inc.

DNA Sequencing by: Greg Lennon, Ph.D.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/www-bio.llnl.gov/bbrp/image/image.html at:

Insert Length: 2019 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham.

Location/Qualifiers

source

1. .325
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1085085"
 /clone_lib="NCI CGAP Gas1"
 /tissue_type="gastric tumor"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Organ: stomach; Vector: Bluescript SK-; Site: 1:
 scori; Site: 2: XhoI; Cloned unidirectionally. Primer:
 Oligo dt. Pooled gastric tumors. 5' adaptor sequence: 5'
 GAATTCGCGCAGG 3' 3' adaptor sequence: 5'
 CTCGAGTTTTTTTTTTT 3' average insert size: 1.0 kb."
 BASE COUNT 55 a 97 c 77 g 96 t
 ORIGIN

Query Match 84.0%; Score 21; DB 9; Length 325;
 Best Local Similarity 100.0%; Pred. No. 0.045;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTTCAGTGGCCGAGATAA 21
 |||||
 Db 32 GGTTCAGTGGCCGAGATAA 12

RESULT 7

T92723/c
 LOCUS T92723 389 bp mRNA linear EST 22-MAR-1995
 DEFINITION Y623c04.s1 Stragatene lung (#937210) Homo sapiens cDNA clone
 IMAGE:118566 3' similar to contains Alu repetitive element; mRNA
 sequence.
 T92723

ACCESSION T92723
 VERSION 1
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 389)

REFERENCE

AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
 Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins,
 M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore,
 B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,
 Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E.,
 Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.
 Generation and analysis of 280,000 human expressed sequence tags
 Genome res. 6 (9), 807-828 (1996)
 97044478

TITLE

CONTACT: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

JOURNAL

COMMENT
 Insert Size: 601
 High quality sequence stops: 359 Source: IMAGE Consortium, LNL
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 601 Std Error: 0.00
 Seq primer: -21ml3
 High quality sequence stop: 359.

Location/Qualifiers

1. .389
 /organism="Homo sapiens"
 /db_xref="CDB:486855"
 /db_xref="taxon:9606"
 /clone="IMAGE:118566"
 /clone_lib="Stragatene lung (#937210)"
 /sex="male"
 /dev_stage="72 years"
 /lab_host="SOLR cells (kanamycin resistant)"
 /note="Organ: lung; Vector: pBluescript SK-; Site: 1: EcoRI
 ; Site: 2: XhoI; Cloned unidirectionally. Primer: Oligo
 dt. normal lung. Average insert size: 1.0 kb; Uni-ZAP XR

FEATURES

source

Query Match 84.0%; Score 21; DB 12; Length 397;
 Best Local Similarity 100.0%; Pred. No. 0.047;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTTCAGTGGCCGAGATAA 21
 |||||
 Db 33 GGTTCAGTGGCCGAGATAA 13

Vector; -5' adaptor sequence: 5' GAATTCGCGCAGG 3' -3'
 adaptor sequence: 5' CTCGAGTTTTTTTTTTT 3"
 BASE COUNT 75 a 113 c 92 g 99 t 10 others
 ORIGIN

Query Match 84.0%; Score 21; DB 14; Length 389;
 Best Local Similarity 100.0%; Pred. No. 0.047;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTTCAGTGGCCGAGATAA 21
 |||||
 Db 57 GGTTCAGTGGCCGAGATAA 37

RESULT 8

BF754298/c
 LOCUS BF754298 397 bp mRNA linear EST 10-JAN-2001
 DEFINITION IL5-CT0519-091000-172-d02 CT0519 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BF754298
 VERSION BF754298.1 GI:12081078
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 397)

REFERENCE

AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663

TITLE

CONTACT: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

JOURNAL

COMMENT
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=IL5&dt=IL5-CT0519-
 091000-172-d02&t3=2000-10-09&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 396.

MEDLINE

Location/Qualifiers

1. .397
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CT0519"
 /dev_stage="Adult"
 /note="Organ: colon; Vector: puc18; Site: 1: SmaI; Site: 2:
 SmaI; A mini-library was made by cloning products derived
 from ORESTES PCR (U.S. Letters Patent application No. 196
 ,716 - Ludwig Institute for Cancer Research) profiles
 into the pUC 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."
 BASE COUNT 76 a 92 c 76 g 153 t
 ORIGIN

FEATURES

source

Query Match 84.0%; Score 21; DB 12; Length 397;
 Best Local Similarity 100.0%; Pred. No. 0.047;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTTCAGTGGCCGAGATAA 21
 |||||
 Db 33 GGTTCAGTGGCCGAGATAA 13

RESULT 9
 AQ384259
 LOCUS
 Definition RPCI11-139H18.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-139H18, DNA sequence.
 ACCESSION AQ384259
 VERSION AQ384259.1 GI:4355282
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 452)
 AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C.
 TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building
 JOURNAL Map Building
 COMMENT Other_GSSs: RPCI11-139H18.TV
 Contact: Shaying Zhao, William Nierman, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbe@tigr.org
 Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (http://bacpac.med.buffalo.edu/ordering) or from http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
 Seq primer: SP6
 Class: BAC ends.
 FEATURES
 source
 Location/Qualifiers
 1..452
 /organism="Homo sapiens"
 /db_xref="GDB:755317"
 /db_xref="taxon:9606"
 /clone="RPCI-11-139H18"
 /clone_lib="RPCI-11"
 /sex="Male"
 /cell_type="Lymphocytes"
 /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"
 BASE COUNT 133 a 98 c 104 g 114 t 3 others
 ORIGIN
 Query Match 84.0%; Score 21; DB 17; Length 452;
 Best Local Similarity 100.0%; Pred. No. 0.049;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTTCGAGTCGAGCGAGATAA 21
 |||||
 Db 322 GGTTCGAGTCGAGCGAGATAA 342
 Query Match 84.0%; Score 21; DB 17; Length 452;
 Best Local Similarity 100.0%; Pred. No. 0.049;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTTCGAGTCGAGCGAGATAA 21
 |||||
 Db 322 GGTTCGAGTCGAGCGAGATAA 342
 RESULT 10
 AQ702820/c
 LOCUS
 Definition HS_5449_B1_H07_07A RPCI-11 Human Male BAC Library Homo sapiens genomic Clone Plate-1025 Col=13 Row=P, DNA sequence.
 ACCESSION AQ702820
 VERSION AQ702820.1 GI:5412246
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 461)
 AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,

Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
 Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
 Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 99380589
 Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering.bac.htm) or from Research Genetics (http://bacpac.med.buffalo.edu/ordering.bac.htm) or from Resear h Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu
 Plate: 1025 row: P column: 13
 Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 461.
 Location/Qualifiers
 1..461
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="Plate-1025 Col=13 Row=P"
 /clone_lib="RPCI-11 Human Male BAC Library"
 /sex="male"
 /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"
 BASE COUNT 103 a 107 c 104 g 144 t 3 others
 ORIGIN
 Query Match 84.0%; Score 21; DB 17; Length 461;
 Best Local Similarity 100.0%; Pred. No. 0.049;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTTCGAGTCGAGCGAGATAA 21
 |||||
 Db 275 GGTTCGAGTCGAGCGAGATAA 255
 RESULT 11
 AW751650
 LOCUS
 Definition IL0-CT0079-030899-106-e11 CT0079 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AW751650
 VERSION AW751650.1 GI:7666582
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 476)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICOR Human Cancer Genome Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?l1=IL06t2-IL0-CT0079-030899-106-ell1t3-1999-08-03&t4=1>)
Seq primer: puc 18 forward
High quality sequence start: 96
High quality sequence stop: 126.

```

nrgn quality sequence stop: 126.
Location/Qualifiers
  1..476
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone_lib="CT0079"
    /dev_stage="Adult"
    /note="Organ: colon; Vector: pUC18; Site_1: SmaI; Site_2:
    SmaI; A mini-library was made by cloning products derived
    from ORESUES PCR (U.S. Letters Patent application No. 196
    7716 - Ludwig Institute for Cancer Research) profiles
    into the pUC 18 vector. Reverse transcription of tissue
    mRNA and cDNA amplification were performed under low
    stringency conditions."
BASE COUNT      180 a      85 c      105 g      106 t
ORIGIN
Query Match      84.0%; Score 21; DB 10; Length 476;
Best Local Similarity 100.0%; Pred. NO. 0.049;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GGTTGCAGTGAGCCGAGATAA  21
        |||||||||||||||||||
Db      391  GGTTGCAGTGAGCCGAGATAA  401

```

RESULT 12
AW265138/c
LOCUS
DEFINITION
481 bp mRNA linear EST 28-DEC-1999
xp80c10.x1 NCI-CGAP_Ov40 Homo sapiens cDNA clone IMAGE:2746674 3'
similar to contains Alu repetitive element,, mRNA sequence.
ACCESSION
AW265138
VERSION
AW265138.1 GI:6641954
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 481)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Andrew Berchuck M.D., John Gillespie M.D.,
Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

```

Possible reversed clone: polyT not found
Seq primer: -40UP from Gibco
High quality sequence stop: 437.
Location/Qualifiers
1. 481
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2745674"
/clone_lib="NCI_CGAP_Ov40"

```

/sex="female"					
/tissue_type="endometrioid ovarian metastasis"					
/lab_host="DH10B"					
/note="Organ: ovary; Vector: pAMP10; cDNA made by oligo-dT priming. Non-directionally cloned into the UDG sites of pAMP10. Size-selected on agarose gel, average insert size 500 bp. Primary library; non-amplified. cDNA Library Preparation: David B. Krizman, Ph.D (NCI). Reference: Krizman et al. (1996) Cancer Research 56:5380-5383."					
95	a	138	c	104	g
					144
					t
BASE COUNT					
ORIGIN					

```

Query Match      84.0%;   Score 21;   DB 10;   Length 481;
Best Local Similarity 100.0%;   Pred. No. 0.049;
Matches 21;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy      1  GGTTCGAGTCAGCCGAGATAA 21
          |||||
Db      141  GGTTCGAGTCAGCCGAGATAA 121

RESULT 13
BE646402/c
LOCUS
DEFINITION
503 bp      mRNA      linear      EST 05-SEP-2000
7886c03.x1  NCI-CGAP-CL11 Homo sapiens cDNA clone IMAGE:3292036 3'
similar to contains Alu repetitive element;contains element PTR5
repetitive element ;, mRNA sequence.
BE646402
BE646402.1  GI:9370713
ACCESION
VERSION
KEYWORDS
SOURCE
human.

```

ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 503)
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap . National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Ash. Alizadeh, John Byrd, M.D., Mike Grever, M.D., Louis M. Staudt, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov Seq primer: -40UP from Gibco High quality sequence stop: 449. Location/Qualifiers 1..503
FEATURES	source

```

FEATURES             SOURCE
  1. 503
    Location/Qualifiers
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone_1="IMAGE:3292036"
      /clone_lib="NCI_CGAP_CL11"
      /tissue_type="B-cell, chronic lymphocytic leukemia"
      /lab_host="DH10B"
      /note="Vector: pT73p-Pac (Pharmacia) with a modified
polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTACCAATCTGAGTGGGAGCGCGCATGCTTTTTTTTTTTTTTTTTTT
T 3']"; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."
  107 a 149 c 121 g 126 t

```

Query Match 84.0%; Score 21; DB 10; Length 503;

Best Local Similarity 100.0%; Pred. No. 0.05; Mismatches 0; Indels 0; Gaps 0;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCGAGTGGCCGAGATAA 21
 Db 75 GGTTCGAGTGGCCGAGATAA 55

RESULT 14
 AQT21912/C
 LOCUS
 DEFINITION
 genomic clone Plate-1137 Col-12 Row-F, DNA sequence.

ACCESSION
 AQT21912
 VERSION
 AQT21912.1 GI:5481581
 KEYWORDS
 GSS.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 (bases 1 to 509)
 Mahairas G.S., Wallace J.C., Smith K., Swartzell S., Holzman T.,
 Keller A., Shaker R., Furlong J., Young J., Zhao S., Adams M.D. and
 Hood L.
 Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

TITLE
 Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome

JOURNAL
 MEDLINE
 COMMENT
 Contact: Mahairas GG, Wallace JC, Hood L
 University of Washington
 High Throughput Sequencing Center
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu

FEATURES
 source
 1..509
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="RPCI-11 Human Male BAC Library"
 /sex="male"
 /note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI;
 Male blood DNA was isolated from one randomly chosen donor
 and partially digested with a combination of EcoRI and
 EcoRI Methylase. Size selected DNA was cloned into the
 pBAC3.6 vector at EcoRI sites"

BASE COUNT
 104 a 123 c 92 g 183 t 7 others

ORIGIN
 Query Match 84.0%; Score 21; DB 17; Length 509;
 Best Local Similarity 100.0%; Pred. No. 0.05;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCGAGTGGCCGAGATAA 21
 Db 123 GGTTCGAGTGGCCGAGATAA 103

RESULT 15
 B37536/C
 LOCUS
 DEFINITION
 sapiens genomic clone Plate-CT 826 Col-21 Row-M, DNA sequence.

ACCESSION
 B37536
 VERSION
 B37536.1 GI:2536905
 KEYWORDS
 GSS.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 (bases 1 to 514)
 Mahairas G.S., Zackrone K.D., Smith T., Tipton S., Schmidt S.,
 Traicoff R., Abajian C., Blanchard A., West A. and Hood L.E.
 Construction of a Characterized Clone Resource for Genomic
 Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
 Tagged Connectors
 Unpublished (1997)
 Contact: Mahairas GG, Zackrone KD, Hood L
 University of Washington
 Seattle, WA 98195, USA
 Tel: (206) 616-8744
 Fax: (206) 685-7301
 Email: kzackrone@u.washington.edu
 Sequence Tagged Connector
 Plate: CT 826 row: M column: 21
 Class: BAC ends
 High quality sequence stop: 514.

TITLE
 Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome

JOURNAL
 MEDLINE
 COMMENT
 Contact: Mahairas GG, Wallace JC, Hood L
 University of Washington
 High Throughput Sequencing Center
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu

FEATURES
 source
 1..514
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CT 826 Col-21 Row-M"
 /clone_lib="CIT Human Genomic Sperm Library C"
 /sex="M"
 /note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
 E-Coli DH10B"

BASE COUNT
 125 a 118 c 91 g 179 t 1 others

ORIGIN
 Query Match 84.0%; Score 21; DB 17; Length 514;
 Best Local Similarity 100.0%; Pred. No. 0.05;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCGAGTGGCCGAGATAA 21
 Db 142 GGTTCGAGTGGCCGAGATAA 122

RESULT 16
 AQT19983
 LOCUS
 DEFINITION
 genomic clone Plate-846 Col-19 Row-A, DNA sequence.

ACCESSION
 AQT19983
 VERSION
 AQT19983.1 GI:5782376
 KEYWORDS
 GSS.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 (bases 1 to 529)
 Mahairas G.S., Wallace J.C., Smith K., Swartzell S., Holzman T.,
 Keller A., Shaker R., Furlong J., Young J., Zhao S., Adams M.D. and
 Hood L.
 Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

TITLE
 Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome

JOURNAL
 MEDLINE
 COMMENT
 Contact: Mahairas GG, Wallace JC, Hood L
 University of Washington
 High Throughput Sequencing Center
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu

FEATURES
 source
 1..529
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="RPCI-11 Human Male BAC Library"
 /sex="male"
 /note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI;
 Male blood DNA was isolated from one randomly chosen donor
 and partially digested with a combination of EcoRI and
 EcoRI Methylase. Size selected DNA was cloned into the
 pBAC3.6 vector at EcoRI sites"

BASE COUNT
 104 a 123 c 92 g 183 t 7 others

ORIGIN
 Query Match 84.0%; Score 21; DB 17; Length 509;
 Best Local Similarity 100.0%; Pred. No. 0.05;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCGAGTGGCCGAGATAA 21
 Db 123 GGTTCGAGTGGCCGAGATAA 103

RESULT 15
 B37536/C
 LOCUS
 DEFINITION
 sapiens genomic clone Plate-CT 826 Col-21 Row-M, DNA sequence.

ACCESSION
 B37536
 VERSION
 B37536.1 GI:2536905
 KEYWORDS
 GSS.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 (bases 1 to 514)
 Mahairas G.S., Zackrone K.D., Smith T., Tipton S., Schmidt S.,
 Traicoff R., Abajian C., Blanchard A., West A. and Hood L.E.
 Construction of a Characterized Clone Resource for Genomic
 Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
 Tagged Connectors
 Unpublished (1997)
 Contact: Mahairas GG, Zackrone KD, Hood L
 University of Washington
 Seattle, WA 98195, USA
 Tel: (206) 616-8744
 Fax: (206) 685-7301
 Email: kzackrone@u.washington.edu
 Sequence Tagged Connector
 Plate: CT 826 row: M column: 21
 Class: BAC ends
 High quality sequence stop: 514.

TITLE
 Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome

JOURNAL
 MEDLINE
 COMMENT
 Contact: Mahairas GG, Wallace JC, Hood L
 University of Washington
 High Throughput Sequencing Center
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu

FEATURES
 source
 1..514
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CT 826 Col-21 Row-M"
 /clone_lib="CIT Human Genomic Sperm Library C"
 /sex="M"
 /note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
 E-Coli DH10B"

BASE COUNT
 125 a 118 c 91 g 179 t 1 others

ORIGIN
 Query Match 84.0%; Score 21; DB 17; Length 514;
 Best Local Similarity 100.0%; Pred. No. 0.05;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCGAGTGGCCGAGATAA 21
 Db 142 GGTTCGAGTGGCCGAGATAA 122

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering.bac.htm>) or from Research Genetics (<http://www.htsc.washington.edu>). BAC end Web Server: <http://www.htsc.washington.edu>
 plate: 846 row: A column: 19
 Seq primer: SP6
 Class: BAC ends
 High quality sequence stop: 529.

FEATURES

source
 Location/Qualifiers
 1..529

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="Plate-846 Col=19 Row=A"
 /clone_lib="RPCI-11 Human Male BAC Library"
 /sex="male"
 /note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
 Male blood DNA was isolated from one randomly chosen donor
 and partially digested with a combination of EcoRI and
 EcoRI Methylase. Size selected DNA was cloned into the
 pBACE3.6 vector at EcoRI sites"

BASE COUNT 146 a 125 c 120 g 133 t 5 others
 ORIGIN

Query Match 84.0%; Score 21; DB 17; Length 529;
 Best Local Similarity 100.0%; Pred. No. 0.05;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGGCCGAGATAA 21
 |||||
 Db 424 GGTTCAGTGGCCGAGATAA 444

RESULT 17
 AQ608587/c

LOCUS HS_5411_B1_L7A RPCI-11 Human Male BAC Library GSS 10-JUN-1999
 DEFINITION genomic clone Plate-987 Col=21 Row=D, DNA sequence.
 ACCESSION AQ608587
 VERSION AQ608587.1 GI:5068581

KEYWORDS GSS.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 539)
 AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
 Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
 Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering.bac.htm>) or from Research Genetics (<http://www.htsc.washington.edu>). BAC end Web Server: <http://www.htsc.washington.edu>
 plate: 987 row: D column: 21
 Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 539.

FEATURES

source
 Location/Qualifiers
 1..539

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="Plate-987 Col=21 Row=D"
 /clone_lib="RPCI-11 Human Male BAC Library"
 /sex="male"
 /note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
 Male blood DNA was isolated from one randomly chosen donor
 and partially digested with a combination of EcoRI and
 EcoRI Methylase. Size selected DNA was cloned into the
 pBACE3.6 vector at EcoRI sites"

BASE COUNT 111 a 142 c 130 g 150 t 6 others
 ORIGIN

Query Match 84.0%; Score 21; DB 17; Length 539;
 Best Local Similarity 100.0%; Pred. No. 0.051;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGGCCGAGATAA 21
 |||||
 Db 158 GGTTCAGTGGCCGAGATAA 138

RESULT 18
 BI494061

LOCUS df107d10.y1 Morton Fetal Cochlea Homo sapiens cDNA clone
 DEFINITION IMAGE:2538690 5', mRNA sequence.
 ACCESSION BI494061
 VERSION BI494061.1 GI:15333405

KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 542)

AUTHORS Robertson,N.G., Khetarpal,U., Gutierrez-Espeleta,G.A., Bieber,F.R.
 and Morton,C.C.
 TITLE Isolation of novel and known genes from a human fetal cochlear cDNA
 library using subtractive hybridization and differential screening
 Genomics 23, 42-50 (1994)
 JOURNAL 95130111
 MEDLINE

COMMENT

Contact: Morton, C. C.
 Departments of Pathology and Obstetrics, Gynecology and
 Reproductive Biology
 Brigham and Women's Hospital
 75 Francis Street, Harvard Medical School, Boston, MA 02115, USA
 Tel: 617 732 7980
 Fax: 617 738 6996
 Email: cmorton@rics.bwh.harvard.edu

DNA sequencing and analyses were performed by National Institutes
 of Health Intramural Sequencing Center (NISC; see
<http://www.nisc.nih.gov>).

This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 PLATE: L2AM6323 row: H column: 19

Seq primer: M13RPI reverse primer (ABI).
 Location/Qualifiers
 1..542

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2538690"
 /clone_lib="Morton Fetal Cochlea"
 /tissue_type="cochlea"
 /dev_stage="16-22 week fetus"

/note="organ: ear; Vector: pBluescript SK-; Site_1: EcoRI;
 Site_2: XhoI; Reference: Genomics 23, 42-50 (1994) Cloned
 unidirectionally. Primer: Oligo dt. Fetal cochlea, normal.
 37% of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb. Uni-ZAP
 XR vector. Library constructed by N. Robertson, C. Morton.
 -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor
 sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"

BASE COUNT 180 a 110 c 126 g 126 t

ORIGIN

Query Match 84.0%; Score 21; DB 13; Length 542;
 Best Local Similarity 100.0%; Pred. No. 0.051;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGGCCGAGATAA 21
 |||||
 Db 479 GGTTCAGTGGCCGAGATAA 499

RESULT 19
 A0708886/c
 LOCUS
 DEFINITION HS_2116_AL_D01_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate-2116 Col-1 Row-G, DNA sequence.

ACCESSION A0708886
 VERSION A0708886.1 GI:5418312
 KEYWORDS GSS.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 570)
 Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
 Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
 Hood,L.L.
 Title: Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999);
 99380589
 High throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones may be purchased from Research Genetics (info@resgen.com).
 BAC end Web Server: http://www.htsc.washington.edu
 Plate: 2116 row: G column: 1
 Seq primer: M13 Reverse
 Class: BAC ends
 High quality sequence stop: 570.

FEATURES

source
 Location/Qualifiers
 1..570
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="Plate-2116 Col-1 Row-G"
 /clone_lib="CIT Approved Human Genomic Sperm Library D"
 /sex="male"
 /note="Organ: sperm; Vector: pBelOAC11; BAC Clones in
 E-Coli DH10B"
 BASE COUNT 145 a 111 c 131 g 177 t 6 others
 ORIGIN

Query Match 84.0%; Score 21; DB 17; Length 570;
 Best Local Similarity 100.0%; Pred. No. 0.051;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGGCCGAGATAA 21
 |||||
 Db 264 GGTTCAGTGGCCGAGATAA 244

RESULT 20
 BI494060/c
 LOCUS
 DEFINITION df107d10.w1 Morton Fetal Cochlea Homo sapiens cDNA clone
 IMAGE:2538690 3', mRNA sequence.

ACCESSION BI494060
 VERSION BI494060.1 GI:15333404
 KEYWORDS EST.

SOURCE

ORGANISM human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 573)
 Robertson,N.G., Khetarpal,U., Gutierrez-Espeleta,G.A., Bieber,F.R.
 and Morton,C.C.
 Title: Isolation of novel and known genes from a human fetal cochlear cDNA
 library using subtractive hybridization and differential screening
 Genomics 23, 42-50 (1994)
 95130111
 JOURNAL
 MEDLINE
 COMMENT
 Contact: Morton, C. C.
 Departments of Pathology and Obstetrics, Synecology and
 Reproductive Biology
 Brigham and Women's Hospital
 75 Francis Street, Harvard Medical School, Boston, MA 02115, USA
 Tel: 617 732 7980
 Fax: 617 738 6996
 Email: cmorton@rics.bwh.harvard.edu
 DNA sequencing and analyses were performed by National Institutes
 of Health Intramural Sequencing Center (NISC; see
 http://www.nisc.nih.gov).
 This clone is available royalty-free through LBNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Plate: L1AM6323 row: H column: 19
 Seq primer: T7 primer.

FEATURES

source
 Location/Qualifiers
 1..573
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2538690"
 /clone_lib="Morton Fetal Cochlea"
 /tissue_type="cochlea"
 /dev_stage="16-22 week fetus"
 /lab_host="SOLR cells (kanamycin resistant)"
 /note="Organ: ear; Vector: pBluescript SK-; Site_1: EcoRI;
 Site_2: XhoI; Reference: Genomics 23, 42-50 (1994) Cloned
 unidirectionally. Primer: Oligo dt. Fetal cochlea, normal.
 37% of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb. Uni-ZAP
 XR Vector. Library constructed by N. Robertson, C. Morton.
 -5' adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor
 sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'."
 BASE COUNT 128 a 126 c 112 g 207 t
 ORIGIN

Query Match 84.0%; Score 21; DB 13; Length 573;
 Best Local Similarity 100.0%; Pred. No. 0.051;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGGCCGAGATAA 21
 |||||
 Db 95 GGTTCAGTGGCCGAGATAA 75

RESULT 21
 A0520887/c
 LOCUS
 DEFINITION HS_5218_B2_B08_T7A RPT-11 Human Male BAC Library Homo sapiens
 genomic clone Plate-794 Col-16 Row-D, DNA sequence.

ACCESSION A0520887
 VERSION A0520887.1 GI:4752012
 KEYWORDS GSS.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 575)
 Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
 Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
 Hood,L.L.
 Title: Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

93380589

MEDLINE
COMMENT

Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu
Plate: 794 row: D column: 16
Seq primer: T7
Class: BAC ends

High quality sequence stop: 575.

FEATURES
source

Location/Qualifiers
1..575
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="RPCI-11 Human Male BAC Library"
/sex="male"
/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"

BASE COUNT
ORIGIN

142 a 131 c 117 g 178 t 7 others

Query Match

Best Local Similarity 84.0%; Score 21; DB 17; Length 575;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGAGCGGAGATAA 21

|||||

Db 126 GTTTCAGTGAGCGGAGATAA 106

RESULT 22

AQ482581

DEFINITION
RPCI-11-243623-TV RPCI-11 Homo sapiens genomic clone RPCI-11-243623
DNA sequence.

ACCESSION

AQ482581

VERSION

AQ482581.1 GI:4669985

KEYWORDS

GSS.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1

Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter

J.C.

Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready

Map Building

Unpublished (1997)

Other_GSSs: RPCI-11-243623.TJ

Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: hbe@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from

Research Genet cs (info@resgen.com). BAC end search page:

FEATURES
source

Location/Qualifiers
1..649
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="RPCI-11-243623"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"

BASE COUNT
ORIGIN

194 a 140 c 144 g 171 t

Query Match

Best Local Similarity 84.0%; Score 21; DB 17; Length 649;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGAGCGGAGATAA 21

|||||

Db 344 GGTTCAGTGAGCGGAGATAA 364

RESULT 23

AQ039496

LOCUS

DEFINITION

CIT-HSP-2327C9-TR CIT-HSP Homo sapiens genomic clone 2327C9, DNA

sequence.

ACCESSION

AQ039496

VERSION

AQ039496.1 GI:3305328

KEYWORDS

GSS.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1

(bases 1 to 666)

Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,

Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,

Simon,M. and Venter,J.C.

Use of a random BAC End Sequence Database for Sequence-Ready Map

Building (1998)

Unpublished (1998)

Other_GSSs: CIT-HSP-2327C9.TF

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC

end search page:

http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.

Seq primer: M3 Reverse

Class: BAC ends.

Location/Qualifiers

1..666

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="2327C9"

/clone_lib="CIT-HSP"

/sex="Male"

/cell_type="Sperm"

/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:

HindIII"

BASE COUNT

209 a 151 c 170 g 136 t

ORIGIN

Query Match

Best Local Similarity 84.0%; Score 21; DB 17; Length 666;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

/clone_lib="NIH_MGC_66"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: ovary; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.8 kb. Library constructed by Life
 Technologies."
 BASE COUNT 280 a 228 c 241 g 165 t
 ORIGIN

Query Match 84.0%; Score 21; DB 12; Length 914;
 Best Local Similarity 100.0%; Pred. No. 0.057;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTTCAGTGCAGCCGAGATAA 21
 |||||
 Db 541 GGTTCAGTGCAGCCGAGATAA 561

RESULT 27
 BF837405
 LOCUS PM2-HT0353-211100-007-f09 HT0353 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BF837405 EST 13-JAN-2001
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 132)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM2&t2=PM2-HT0353-
 211100-007-f09&t3=2000-11-21&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 3
 High quality sequence stop: 131.
 Location/Qualifiers
 1. 132
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0353"
 /dev_stage="Adult"
 /note="Organ: head, neck; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 54 a 30 c 34 g 14 t
 ORIGIN

Query Match 80.0%; Score 20; DB 12; Length 132;
 Best Local Similarity 100.0%; Pred. No. 0.14;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTTCAGTGCAGCCGAGATA 20
 |||||
 Db 68 GGTTCAGTGCAGCCGAGATA 87

RESULT 29
 AA709114/c
 LOCUS
 Query Match 80.0%; Score 20; DB 12; Length 139;
 Best Local Similarity 100.0%; Pred. No. 0.14;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTTCAGTGCAGCCGAGATA 20
 |||||
 Db 68 GGTTCAGTGCAGCCGAGATA 87

RESULT 28
 BF878448
 LOCUS MRO-ET0109-191100-002-c07 ET0109 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BF878448
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 139)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR0&t2=MR0-ET0109-
 191100-002-c07&t3=2000-11-19&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 46
 High quality sequence stop: 138.
 Location/Qualifiers
 1. 139
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="ET0109"
 /dev_stage="Adult"
 /note="Organ: lung, tumor; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 31 a 33 c 48 g 27 t
 ORIGIN

Query Match 80.0%; Score 20; DB 12; Length 139;
 Best Local Similarity 100.0%; Pred. No. 0.14;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTTCAGTGCAGCCGAGATA 20
 |||||
 Db 68 GGTTCAGTGCAGCCGAGATA 87

RESULT 29
 AA709114/c
 LOCUS
 Query Match 80.0%; Score 20; DB 12; Length 139;
 Best Local Similarity 100.0%; Pred. No. 0.14;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTTCAGTGCAGCCGAGATA 20
 |||||
 Db 68 GGTTCAGTGCAGCCGAGATA 87

RESULT 28
 BF878448
 LOCUS MRO-ET0109-191100-002-c07 ET0109 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BF878448
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 139)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR0&t2=MR0-ET0109-
 191100-002-c07&t3=2000-11-19&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 46
 High quality sequence stop: 138.
 Location/Qualifiers
 1. 139
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="ET0109"
 /dev_stage="Adult"
 /note="Organ: lung, tumor; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 31 a 33 c 48 g 27 t
 ORIGIN

Putative full length read
The vector to vector length is 186
Insert Length: 533 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham.

FEATURES

source
1. .173
/organism="Homo sapiens"
/db_xref="GDB:1285026"
/db_xref="taxon:9606"
/clone="IMAGE:376770"
/clone_lib="Soares_fetal_heart_NBHH19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: heart; Vector: pT73D (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAGTGGAGCGCGCATCTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NBHH19W."

BASE COUNT 51 a 36 c 52 g 31 t 3 others
ORIGIN

Query Match 80.0%; Score 20; DB 9; Length 173;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGGAGCGGAGATA 20
|||||
Db 120 GGTTCAGTGGAGCGGAGATA 139

RESULT 32
B73330/c

LOCUS B73330 175 bp DNA linear GSS 21-JUN-1998
DEFINITION CIT-HSP-712K9-TV CIT-HSP Homo sapiens genomic clone 712K9, DNA sequence.

ACCESSION B73330
VERSION B73330.1 GI:2712481
KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Kim,U.-J., Adams,M.D. and Simon,M.I.
TITLE Determination of clone end sequences of human Bacterial Artificial Chromosomes

JOURNAL Unpublished (1997)
COMMENT Other_GSSs: CIT-HSP-712K9.TP
Contact: Ung-Jin Kim
Caltech Genome Research Lab
California Institute of Technology
Division of Biology, MS 147-75, Pasadena, CA 91125, USA
Tel: 626 796 7066
Fax: 626 395 4901
Email: ung@ash.tree.caltech.edu

Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: T7
Class: BAC ends.

FEATURES

source
1. .175
/organism="Homo sapiens"
/db_xref="GDB:550829g"
/db_xref="taxon:9606"

/clone="712K9"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2: HindIII"

BASE COUNT 36 a 47 c 28 g 64 t
ORIGIN

Query Match 80.0%; Score 20; DB 17; Length 175;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGGAGCGGAGATA 20
|||||
Db 158 GGTTCAGTGGAGCGGAGATA 139

RESULT 33
AA348526/c

LOCUS AA348526 184 bp mRNA linear EST 21-APR-1997
DEFINITION ESP55073 Hippocampus II Homo sapiens cDNA 5' end similar to EST containing Alu repeat, mRNA sequence.

ACCESSION AA348526
VERSION AA348526.1 GI:2000835
KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Adams,M.D., Soares,M.B., Kerlavage,A.R., Fields,C. and Venter,J.C.
TITLE Rapid cDNA sequencing (expressed sequence tags) from a directionally cloned human infant brain cDNA library

JOURNAL Nat. Genet. 4, 373-380 (1993)
MEDLINE 94004965
COMMENT Other_ESTs: EST55072
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlavet@tigr.org

For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M3 Reverse.

FEATURES

source
1. .184
/organism="Homo sapiens"
/db_xref="ATCC (inhost):149782"
/db_xref="taxon:9606"
/clone_lib="Hippocampus II"
/sex="male"
/tissue_type="hippocampus"
/dev_stage="adult, 74 yrs"
/note="Organ: brain; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 39 a 51 c 57 g 37 t
ORIGIN

Query Match 80.0%; Score 20; DB 9; Length 184;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGGAGCGGAGATA 20
|||||
Db 65 GGTTCAGTGGAGCGGAGATA 46

RESULT 34
BQ251032/c

LOCUS BQ251032 185 bp mRNA linear EST 03-MAY-2002
 DEFINITION TAE25034A06R TAE25 Triticum aestivum cDNA clone TAE25034A06R, mRNA sequence.
 ACCESSION BQ251032
 VERSION BQ251032.1 GI:20446908
 KEYWORDS EST.
 SOURCE bread wheat.
 ORGANISM Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.
 REFERENCE 1 (bases 1 to 185)
 AUTHORS Cloutier,S.
 TITLE Wheat functional genomics - Glenlea developing seeds cDNA libraries
 JOURNAL Unpublished (2002)
 COMMENT Contact: Dr. Sylvie Cloutier
 Cereal Research Centre, Agriculture and Agri-food Canada
 195 Daffoe Rd, Winnipeg, MB, Canada R3T 2M9
 Tel: (204) 983-2340
 Fax: (204) 983-4604
 Email: scloutier@em.agr.ca
 was cloned directionally, not all sequences generated with reverse primer were from the 5' end (same with forward primer and 3' end).
 Average insert size is >870 bp
 Plate: 034 row: A column: 06
 Seq primer: M13 Reverse

FEATURES
 source Location/Qualifiers
 1..185

/organism="Triticum aestivum"
 /cultivar="Glenlea"
 /db_xref="taxon:4565"
 /clone="TAE25034A06R"
 /clone_lib="TAE25"
 /tissue_type="developing seeds"
 /dev_stage="25 days after anthesis"
 /lab_host="E. coli DH10B"
 /note="Vector: pCMV-SPORT6.0 (Invitrogen Technologies); Site_1: NotI; Site_2: XbaI; mRNA obtained from wheat seeds of cultivar Glenlea 25 days post-anthesis"

BASE COUNT 37 a 47 c 43 g 58 t
 ORIGIN
 |||||

Query Match 80.0%; Score 20; DB 14; Length 185;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGCAGCGAGATA 20
 |||||
 Db 72 GGTTCAGTGCAGCGAGATA 53

RESULT 35
 H02828/c
 LOCUS

DEFINITION YJ43a04.s1 Soares placenta Nb2Hp Homo sapiens cDNA clone IMAGE:151470 3' similar to contains Alu repetitive element; contains XTR repetitive element;; mRNA sequence.

ACCESSION H02828
 VERSION H02828.1 GI:865761
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 189)
 AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.

TITLE The WashU-Merck EST Project
 JOURNAL Unpublished (1995)
 COMMENT Contact: Wilson RK
 Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Insert Size: 293
 Source: IMAGE Consortium, LLNL
 this clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 293 Std Error: 0.00
 Seq primer: Promega -21ml3
 High quality sequence stop: 374.

FEATURES
 source Location/Qualifiers
 1..189

/organism="Homo sapiens"
 /db_xref="GDB:563571"
 /db_xref="taxon:9606"
 /clone="IMAGE:151470"
 /clone_lib="Soares placenta Nb2Hp"
 /sex="Female"
 /dev_stage="placenta obtained at birth (full term)"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: placenta; Vector: p773D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAAGATTCGGCGCGAGGATTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 35 a 54 c 44 g 56 t
 ORIGIN
 |||||

Query Match 80.0%; Score 20; DB 14; Length 189;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGCAGCGAGATA 20
 |||||
 Db 65 GGTTCAGTGCAGCGAGATA 46

RESULT 36
 R44194/c
 LOCUS

DEFINITION R44194 191 bp mRNA linear EST 22-MAY-1995 YG34D02.s1 Soares infant brain INIB Homo sapiens cDNA clone IMAGE:34561 3' similar to contains Alu repetitive element;; mRNA sequence.

ACCESSION R44194
 VERSION R44194.1 GI:822058
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 191)
 AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.

TITLE The WashU-Merck EST Project
 JOURNAL Unpublished (1995)
 COMMENT On May 9, 1995 this sequence version replaced gi:802918.
 Contact: Wilson RK
 Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Insert Size: 1200

High quality sequence stops: 178 Source: IMAGE Consortium, LLNL
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 1200 Std Error: 0.00

Seq primer: Promega -2lm13

High quality sequence stop: 178.

Location/Qualifiers

1. .191

/organism="Homo sapiens"

/db_xref="GDB:406908"

/db_xref="taxon:9606"

/clone="IMAGE:34561"

/clone_lib="Soares infant brain LNIB"

/sex="female"

/dev_stage="73 days post natal"

/lab_host="DH10B (ampicillin resistant)"

/note="Organ: whole brain; Vector: Lfamid BA; Site.1: Not

I; Site.2: Hind III; 1st strand cDNA was primed with a Not

I - oligo(dT) primer [5'

AACTGGAAGATTCCGCCGCGAGGAATTTTTTTTTTTT 3'];

double-stranded cDNA was ligated to Hind III adaptors

(Pharmacia), digested with Not I and directionally cloned

into the Not I and Hind III sites of the Lfamid BA vector.

Library went through one round of normalization. Library

constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 46 a 53 c 46 g 44 t 2 others

ORIGIN

Query Match 80.0%; Score 20; DB 14; Length 191;

Best Local Similarity 100.0%; Pred. No. 0.15;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGGCGGAGATA 20

Db 75 GGTTCAGTGGCGGAGATA 56

|||||

RESULT 37

LOCUS AI241886/c

DEFINITION 194 bp mRNA linear EST 04-NOV-1998

qu70g06.x1 NCI CGAP_Brn35 Homo sapiens cDNA clone IMAGE:1977466 3'

similar to contains Alu repetitive element; mRNA sequence.

AI241886

AI241886

AI241886.1 GI:3837283

EST.

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 194)

NCI/NINDS-CoAP http://www.ncbi.nlm.nih.gov/ncicqap.

National Cancer Institute / National Institute of Neurological

Disorders and Stroke, Brain Tumor Genome Anatomy Project

(CGAP/BTGAP), Tumor Gene Index

Unpublished (1998)

Contact: Robert Strausberg, Ph.D.

Email: cgaps-r@mail.nih.gov

unknown library type

Seq primer: -400P from Gibco

High quality sequence stop: 183.

Location/Qualifiers

1. .194

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1977466"

/clone_lib="NCI CGAP_Brn35"

/tissue_type="tumor, 5 pooled (see description)"

/lab_host="DH10B"

/note="Organ: brain; Vector: pCMV-SPORT6; Site.1: SalI;

Site.2: NotI; Cloned unidirectionally. Primer: oligo dt.

Average insert size 1.33 kb. Tumor types include:

meningioma, oligodendroglioma, astrocytoma (grade II),

medulloblastoma, astrocytoma (grade IV). Life Technologies

catalog #: 11544-012"

36 a 55 c 52 g 51 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 20; DB 9; Length 194;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGGCGGAGATA 20

Db 65 GGTTCAGTGGCGGAGATA 46

|||||

RESULT 38

LOCUS H03719

DEFINITION 201 bp mRNA linear EST 20-JUN-1995

Y43a04.r1 Soares placenta Nb2HP Homo sapiens cDNA clone

IMAGE:151470 5' similar to contains Alu repetitive element; contains

XTR repetitive element ; mRNA sequence.

H03719

H03719.1 GI:866652

EST.

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 201)

Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M., Holman

M., Hultman M., Kucaba T., Le M., Lennon G., Marra N., Parsons J.,

Rifkin L., Rohlfing J., Soares M., Tan F., Trevaskis E., Waterston

R., Williamson A., Wohldmann P. and Wilson R.

The WashU-Merck EST Project

Unpublished (1995)

Contact: Willson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Insert Size: 293

Source: IMAGE Consortium, LLNL

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 293 Std Error: 0.00

Seq primer: M13RP1

High quality sequence stop: 321.

Location/Qualifiers

1. .201

/organism="Homo sapiens"

/db_xref="GDB:563571"

/db_xref="taxon:9606"

/clone="IMAGE:151470"

/clone_lib="Soares placenta Nb2HP"

/sex="Female"

/dev_stage="placenta obtained at birth (full term)"

/lab_host="DH10B (ampicillin resistant)"

/note="Organ: placenta; Vector: pT73D (Pharmacia) with a

modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st

strand cDNA was primed with a Not I - oligo(dT) primer [5'

AACTGGAGAAATTCGCCGCGAGGAATTTTTTTTTTTT 3'];

double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT73 vector. Library

went through one round of normalization. Library

constructed by Bento Soares and M.Fatima Bonaldo.

60 a 44 c 61 g 35 t

BASE COUNT

ORIGIN

Query Match 80.0%; Score 20; DB 14; Length 201;

Best Local Similarity 100.0%; Pred. No. 0.15;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGGCGGAGATA 20

Db 151 GGTTCAGTGGCGGAGATA 170

|||||

RESULT 39
 N73648
 LOCUS
 DEFINITION
 211 bp mRNA linear EST 19-MAR-1996
 clone IMAGE:289188 3' similar to contains Alu repetitive element;;
 mRNA sequence.
 ACCESSION
 N73648
 VERSION
 N73648.1 GI:1230933
 KEYWORDS
 EST.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 211)
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman,
 M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
 Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisan, E., Waterston,
 R., Williamson, A., Wohlmann, P. and Wilson, R.
 TITLE
 The WashU-Merck EST Project
 JOURNAL
 Unpublished (1995)
 COMMENT
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: ml3 -40 forward
 High quality sequence stop: 202.
 FEATURES
 Location/Qualifiers
 source
 1..211
 /organism="Homo sapiens"
 /db_xref="GDB:3905028"
 /db_xref="taxon:9606"
 /clone="IMAGE:289188"
 /clone_lib="Soares_multiple_sclerosis_2NdbMSP"
 /sex="male"
 /tissue_type="multiple sclerosis lesions"
 /dev_stage="Age 46"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Vector: p773b (Pharmacia) with a modified
 polylinker V-TYPE: Phagemid; Site_1: Not I; Site_2: Eco RI
 ; 1st strand cDNA was primed with a Not I - oligo(dT)
 primer [5'
 TGTACCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTTTTTTTTTT 3']
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified p773 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M. Fatima Bonaldo. RNA from 4 multiple sclerosis
 lesions from one patient was kindly provided by Dr. Kevin
 G. Becker (NINDS/NIH)."

BASE COUNT 45 a 64 c 59 g 43 t
 ORIGIN
 Query Match 80.0%; Score 20; DB 14; Length 211;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTTCAGTCAGCCGAGATA 20
 ||||||||||||||||||
 Db 148 GGTTCAGTCAGCCGAGATA 167

RESULT 40
 N772536
 LOCUS
 DEFINITION
 212 bp mRNA linear EST 04-MAY-2000
 clone IMAGE:2692720 5' similar to contains Alu repetitive element; contains element MER6
 repetitive element ; mRNA sequence.
 ACCESSION
 N772536
 VERSION
 N772536.1 GI:7704692
 KEYWORDS
 EST.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 212)
 NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: ccgaps-remail.nih.gov
 Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui, M.D.,
 Michael Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: David B. Krizman, Ph.D.
 cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCGAP clone distribution information can be

ACCESSION
 N772536
 VERSION
 N772536.1 GI:7704692
 KEYWORDS
 EST.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 212)
 NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: ccgaps-remail.nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
 R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL, send email to:
 info@image.llnl.gov
 Seq primer: -40RP from Gibco
 High quality sequence stop: 82.
 FEATURES
 Location/Qualifiers
 source
 1..212
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2692720"
 /clone_lib="NCI-CCGAP_Li5"
 /tissue_type="hepatic adenoma"
 /lab_host="DH10B"
 /note="Organ: liver; Vector: pCMV-SPORT4; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 0.8 Kb."
 BASE COUNT 120 a 28 c 44 g 20 t
 ORIGIN
 Query Match 80.0%; Score 20; DB 10; Length 212;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTTCAGTCAGCCGAGATA 20
 ||||||||||||||||||
 Db 34 GGTTCAGTCAGCCGAGATA 53

RESULT 41
 N772536
 LOCUS
 DEFINITION
 231 bp mRNA linear EST 21-AUG-1997
 clone IMAGE:1010380
 similar to contains Alu repetitive element; contains element XTR
 repetitive element ; mRNA sequence.
 ACCESSION
 N772536
 VERSION
 N772536.1 GI:1849833
 KEYWORDS
 EST.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 231)
 NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: ccgaps-remail.nih.gov
 Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui, M.D.,
 Michael Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: David B. Krizman, Ph.D.
 cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Insert length: 521 Std Error: 0.00
 Seq primer: -28m13 rev1 ET from Amer-sham
 High quality sequence stop: 225.

FEATURES

source

Location/Qualifiers
 1..231
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1010380"
 /sex="Male"
 /dev_stage="45 years old"
 /lab_host="DH10B"
 /note="Vector: pAMP10; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA was primed with oligo(dT)17 on 50 ng of DNase-treated, total cellular RNA obtained from 5,000-10,000 microdissected preneoplastic cells histologically-determined to be prostatic intraepithelial neoplasia 2 (PIN2) cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into pAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."
 BASE COUNT 43 a 65 c 55 g 68 t
 ORIGIN

Query Match 80.0%; Score 20; DB 9; Length 231;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATA 20
 |||||
 Db 67 GTTGCAGTGAGCCGAGATA 48

RESULT 42

BE261473

LOCUS

BE261473 233 bp mRNA linear EST 26-OCT-2000
 601148613f1 NTH_MGC_19 Homo sapiens cDNA clone IMAGE:3163875 5',
 mRNA sequence.

ACCESSION

BE261473

VERSION

BE261473.1 GI:9133661

KEYWORDS

EST

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 233)

NIH-MGC <http://mgs.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LCM125 row: j column: 04

High quality sequence stop: 233.

Location/Qualifiers

1..233

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3163875"

/clone_lib="NIH_MGC_19"

/tissue_type="neuroblastoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:

EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

BASE COUNT 59 a 64 c 70 g 40 t
 ORIGIN

Query Match

Best Local Similarity 100.0%; Score 20; DB 10; Length 233;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATA 20

|||||

Db 165 GTTGCAGTGAGCCGAGATA 184

RESULT 43

BF962750

LOCUS

BF962750 234 bp mRNA linear EST 22-JAN-2001

QV2-NN0045-181200-560-d05 NN0045 Homo sapiens cDNA, mRNA sequence.

ACCESSION

BF962750

VERSION

BF962750.1 GI:12380025

KEYWORDS

EST..

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 234)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R.,

Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F.,

Goldman, G. H., Carvalho, A. F., Matsukuma, A., Baia, G. S., Simpson, D. H.,

Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare,

M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and

Simpson, A. J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

Contact: Simpson A. J. G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SF,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV2&t2=QV2-NN0045-181200-560-d05&t3=2000-12-18&t4=1>

Seq primer: puc 18 forward

High quality sequence stop: 234.

Location/Qualifiers

1..234

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="NN0045"

/dev_stage="Adult"

/note="Organ: nervous_normal; Vector: puc18; Site_1: SmaI;

Site_2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

BASE COUNT 87 a 39 c 62 g 46 t
 ORIGIN

Query Match

Best Local Similarity 100.0%; Score 20; DB 12; Length 234;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGGAGCCGAGATA 20
 |||
 Db 44 GGTTCAGTGGAGCCGAGATA 63

RESULT 44
 AA380519/C
 LOCUS
 DEFINITION EST93513 Supt cells Homo sapiens cDNA 5' end similar to EST
 ACCESSION AA380519
 VERSION AA380519.1 GI:2032859
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 237)
 AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
 ,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
 ,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,
 Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
 ,L.M., Fitzhugh,W.M., Fritchman,J.D., Geoghagen,N.S., Glodek,A.,
 Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M.,
 Kelley,J.C., Liu,L.-L., Marmaros,S.M., Merrick,J.M.,
 Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
 Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
 Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
 Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W.,
 Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
 Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L.,
 Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon
 ,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
 Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns
 JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
 MEDLINE 96026280
 COMMENT

Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423

Email: arkerlav@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seq primer: M13 Reverse.

FEATURES
 source
 Location/Qualifiers

1..237
 /organism="Homo sapiens"
 /db_xref="ATCC (inhost):184934"
 /db_xref="taxon:9606"
 /clone_lib="Supt cells"
 /cell_type="T-lymphocyte"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI"

BASE COUNT 44 a 69 c 55 g 69 t
 ORIGIN

Query Match 80.0%; Score 20; DB 9; Length 237;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGGAGCCGAGATA 20
 |||
 Db 119 GGTTCAGTGGAGCCGAGATA 100

RESULT 45
 B87492

LOCUS
 DEFINITION RPC111-29B18.TV RPC1-11 Homo sapiens genomic clone RPC1-11-29B18,
 DNA sequence.

ACCESSION B87492
 VERSION B87492.1 GI:2928624
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 239)
 AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
 Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
 Venter,J.C.

USE OF BAC End Sequences for Sequence-Ready Map Building (1998)
 Unpublished (1998)
 CONTACT: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdadams@tigr.org

Clones are derived from the human BAC library RPC1-11. For BAC
 library availability, please contact Pieter de Jong
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from
 Research Genetics (<http://resgen.com>). BAC end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
 Seq primer: T7

Class: BAC ends.
 Location/Qualifiers
 1..239
 /organism="Homo sapiens"
 /db_xref="GDB:7510793"
 /db_xref="taxon:9606"
 /clone="RPC1-11-29B18"
 /clone_lib="RPC1-11"
 /sex="Male"
 /cell_type="Lymphocytes"
 /note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
 RPC111 Human Male BAC Library"

FEATURES
 source

80 a 50 c 57 g 52 t
 BASE COUNT
 ORIGIN

Query Match 80.0%; Score 20; DB 17; Length 239;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGGAGCCGAGATA 20
 |||
 Db 62 GGTTCAGTGGAGCCGAGATA 81

Search completed: November 24, 2002, 15:01:44
 Job time: 1554.5 secs

GenCore version 5.1.3
 Copyright (c) 1993 - 2002 CompuGen Ltd.
 OM nucleic - nucleic search, using sw model
 Run on: November 24, 2002, 12:19:18 ; Search time 211.5 Seconds
 (without alignments)
 266.194 Million cell updates/sec
 Title: US-09-784-423-124
 Perfect score: 25
 Sequence: 1 GGTTCAGTGCAGCCGATAGAGT 25
 Scoring table: OLIGO_NUC
 Gapop 60.0 , Gapext 60.0
 Searched: 2185239 seqs, 112599159 residues
 Word size : 0
 Total number of hits satisfying chosen parameters: 4370478
 Minimum DB seq length: 0
 Maximum DB seq length: 200000000
 Post-processing: Listing first 45 summaries

Database : N.Geneseq_101002.*
 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
 24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
 Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	100.0	25	20	AAZ27824
2	25	100.0	1000	20	AAZ27732
3	21	84.0	325	22	ABAI7408
4	21	84.0	355	22	AAF65697
5	21	84.0	608	24	ABN65341
6	21	84.0	946	22	AAK70967
7	21	84.0	1242	22	AAH48569
8	21	84.0	1714	24	ABA92605
9	21	84.0	2150	23	ABV22677

c 10	10	84.0	2150	23	ABV28503	Human prostate exp
c 11	21	84.0	2843	22	AAH48582	Human fascin DNA f
c 12	21	84.0	5537	22	AAK84869	Human immune/haema
c 13	21	84.0	11674	22	AAK70965	Human immune/haema
c 14	21	84.0	13055	22	AAH48621	Human fascin DNA f
c 15	21	84.0	16951	22	AAH48620	Human fascin DNA f
c 16	21	84.0	16951	22	AAH48622	Human fascin DNA f
c 17	21	84.0	19286	22	AAH48622	Human fascin DNA f
c 18	21	84.0	19286	22	AAH48622	Human fascin DNA f
c 19	21	84.0	28149	22	ABAI7349	Human testicular a
c 20	21	84.0	30360	22	ABAI7361	Human nervous syst
c 21	21	84.0	32192	22	AAK62759	Human immune/haema
c 22	21	84.0	38258	24	AAH17963	Human nervous syst
c 23	21	84.0	49999	20	AAH23903	DNA encoding Colon
c 24	20	80.0	95	21	AAH15906	Human LOBO homolog
c 25	20	80.0	171	22	AAH81714	Human secreted pro
c 26	20	80.0	172	22	AAH81714	Human immune/haema
c 27	20	80.0	180	21	AAH83823	Human secreted pro
c 28	20	80.0	193	21	AAH11950	Human secreted pro
c 29	20	80.0	205	24	AAH04572	Human secreted pro
c 30	20	80.0	236	21	AAH16895	Activated T-cell d
c 31	20	80.0	304	21	AAH04849	Human secreted pro
c 32	20	80.0	320	22	AAH68890	Human secreted pro
c 33	20	80.0	329	22	AAH70549	Human cervical can
c 34	20	80.0	384	24	ABL79179	Human polynucleoti
c 35	20	80.0	388	22	AAH2847	Human ovarian can
c 36	20	80.0	425	23	ABV16660	Human polynucleoti
c 37	20	80.0	428	23	AAH3705	Human prostate exp
c 38	20	80.0	436	23	ABV49562	Human polynucleoti
c 39	20	80.0	436	23	ABV49562	Human prostate exp
c 40	20	80.0	451	22	AAH62995	Human prostate exp
c 41	20	80.0	451	22	AAH62995	Human immune/haema
c 42	20	80.0	459	23	ABV45555	Human prostate exp
c 43	20	80.0	472	23	ABV47512	Human prostate exp
c 44	20	80.0	485	23	ABV46458	Human prostate exp
c 45	20	80.0	533	23	ABV42598	Human prostate exp

ALIGNMENTS

RESULT 1
 ID AAZ27824 standard; DNA; 25 BP.
 AC AAZ27824;
 DT 23-DEC-1999 (first entry)
 DE PCR primer for human DNA marker clone S132.
 XX Tandem repeat sequence; DNA isolation; intermediate tandem repeat;
 XX IIR sequence; pentanucleotide tandem repeat; stutter artifact;
 KW DNA typing; DNA profiling; linkage analysis; criminal justice;
 KW paternity testing; animal lineage analysis; microsatellite loci;
 KW polymorphism detection; PCR primer; ss.
 OS Synthetic.
 OS Homo sapiens.
 XX WO9940194-A1.
 XX 12-AUG-1999.
 PD 04-FEB-1999; 99WO-US02345.
 XX 04-FEB-1998; 98US-0018584.
 XX (PROM-) PROMEGA CORP.
 XX Schumm JW, Bachner JW;
 XX WPI; 1999-590696/50.

XX Isolating DNA containing intermediate tandem repeat sequences, useful
 PT in DNA profiling
 XX
 XX Claim 30: Page 21; 111pp; English.
 XX
 CC This sequence is a PCR primer for a human DNA marker clone used in the
 CC method of the invention. The method is for isolating a fragment of DNA
 CC containing an intermediate tandem repeat (ITR) sequence using
 CC hybridization selection, and comprises: (a) providing several DNA
 CC fragments, at least one of which contains an ITR sequence, a region of
 CC the DNA fragment which contains at least one repeat unit consisting of a
 CC sequence of five, six or seven bases repeated in tandem at least two
 CC times; (b) providing a stationary support having at least one
 CC oligonucleotide associated with it, where the oligonucleotide includes a
 CC sequence of nucleotides which is complementary to a portion of the ITR
 CC sequence; and (c) combining the DNA fragments with the support under
 CC conditions where the DNA fragments including the DNA fragment containing
 CC the ITR sequence hybridize to the support. The method is particularly
 CC used to isolate DNA containing pentanucleotide tandem repeat sequences as
 CC well as to detect target ITR DNA sequences having a low incidence of
 CC stutter artifacts (no more than 2.4%). The method is useful in DNA
 CC profiling for linkage analysis, criminal justice, paternity testing and
 CC other forensic and medical uses. DNA typing is also useful for confirming
 CC the lineage of horses, dogs and other prize animals. The invention
 CC overcomes problems related to the use of microsatellite loci in DNA
 CC profiling. The method can detect polymorphisms with a low incidence of
 CC stutter artifacts, which has previously been a problem in interpreting
 CC allelic content of loci. The development of markers based on larger
 CC repeat units, enables easier separation of the fragments on
 CC electrophoretic gels. This allows the simultaneous analysis of more loci.
 XX
 SQ Sequence 25 BP; 7 A; 3 C; 10 G; 5 T; 0 other;

Query Match 100.0%; Score 25; DB 20; Length 25;
 Best Local Similarity 100.0%; Pred. No. 2.5e-05;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCGAGTGAGCCGAGATAGAGT 25
 |||||
 Db 1 GGTTCGAGTGAGCCGAGATAGAGT 25

RESULT 2

ID AA227732 standard; DNA; 1000 BP.

XX AC AA227732;

XX DT 23-DEC-1999 (first entry)

XX DE Human DNA marker clone s132.

XX KW Tandem repeat sequence; DNA isolation; intermediate tandem repeat;
 KW ITR sequence; pentanucleotide tandem repeat; stutter artifact;
 KW DNA typing; DNA profiling; linkage analysis; criminal justice;
 KW paternity testing; animal lineage analysis; microsatellite loci;
 KW polymorphism detection; ds.

XX OS Homo sapiens.

XX PN WC9940194-A1.

XX PD 12-AUG-1999.

XX PF 04-FEB-1999; 99WO-US02345.

XX PR 04-FEB-1998; 98US-0018584.

XX PA (PROM-) PROMEGA CORP.

XX PI Schumm JW, Bacher JW;

XX

DR WPI; 1999-590696/50.
 XX Isolating DNA containing intermediate tandem repeat sequences, useful
 PT in DNA profiling
 XX
 XX Claim 22: Page 78; 111pp; English.
 XX
 CC This sequence represents a human DNA marker clone used in the method of
 CC the invention. The method is for isolating a fragment of DNA containing
 CC an intermediate tandem repeat (ITR) sequence using hybridization
 CC selection, and comprises: (a) providing several DNA fragments, at least
 CC one of which contains an ITR sequence, a region of the DNA fragment which
 CC contains at least one repeat unit consisting of a sequence of five, six
 CC or seven bases repeated in tandem at least two times; (b) providing a
 CC stationary support having at least one oligonucleotide associated with
 CC it, where the oligonucleotide includes a sequence of nucleotides which is
 CC complementary to a portion of the ITR sequence; and (c) combining the DNA
 CC fragments with the support under conditions where the DNA fragments
 CC including the DNA fragment containing the ITR sequence hybridize to the
 CC support. The method is particularly used to isolate DNA containing
 CC pentanucleotide tandem repeat sequences as well as to detect target ITR
 CC DNA sequences having a low incidence of stutter artifacts (no more than
 CC 2.4%). The method is useful in DNA profiling for linkage analysis,
 CC criminal justice, paternity testing and other forensic and medical uses.
 CC DNA typing is also useful for confirming the lineage of horses, dogs and
 CC other prize animals. The invention overcomes problems related to the use
 CC of microsatellite loci in DNA profiling. The method can detect
 CC polymorphisms with a low incidence of stutter artifacts, which has
 CC previously been a problem in interpreting allelic content of loci. The
 CC development of markers based on larger repeat units, enables easier
 CC separation of the fragments on electrophoretic gels. This allows the
 CC simultaneous analysis of more loci.
 XX
 SQ Sequence 1000 BP; 302 A; 240 C; 231 G; 226 T; 1 other;

Query Match 100.0%; Score 25; DB 20; Length 1000;

Best Local Similarity 100.0%; Pred. No. 2.4e-05;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCGAGTGAGCCGAGATAGAGT 25

|||||

Db 441 GGTTCGAGTGAGCCGAGATAGAGT 465

RESULT 3

ABAL7408/c

ID ABAL7408 standard; DNA; 325 BP.

XX AC ABAL7408;

XX DT 23-JAN-2002 (first entry)

XX DE Human nervous system related polynucleotide SEQ ID NO 9739.

XX KW Human; neotropic; neuroprotective; cytostatic; dermatological; virucide;
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnary;
 KW antiparkinsonian; antiskilling; antianaemic; antiarthritis; cancer;
 KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
 KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
 KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.

XX OS Homo sapiens.

XX PN WC0200159063-A2.

XX PD 16-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01334.

XX PR 31-JAN-2000; 2000US-0179065.

XX PR 04-FEB-2000; 2000US-0180628.

XX PR 24-FEB-2000; 2000US-0184664.

CC The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (AB114578-AB18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 325 BP; 55 A; 97 C; 77 G; 96 T; 0 other;
Query Match 84.0%; Score 21; DB 22; Length 325;
Best Local Similarity 100.0%; Pred. No. 0.0052;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTCAGCCGAGATAA 21
D5 32 GGTTCAGTCAGCCGAGATAA 12
|||||

RESULT 4
AAF65697/c
ID AAF65697 standard; cDNA; 355 BP.

AC AAF65697;

DT 09-APR-2001 (first entry)

DE Novel human polynucleotide, SEQ ID NO: 1453.

KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
breast cancer; lung cancer; cancer detection; ss.

OS Homo sapiens.

XX WO200102568-A2.

XX 11-JAN-2001.

XX 30-JUN-2000; 2000WO-US18374.

XX 02-JUL-1999; 99US-0142310.

XX 02-JUL-1999; 99US-0142311.

PA (CHIR) CHIRON CORP.

PA (HYSE-) HYSEQ INC.

XX Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Criktenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LW, Strache-Crain B;
XX WPI; 2001-091805/10.

XX Library of polynucleotides for diagnosing a cancerous state of a

PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX Claim 9; Page 752; 1046pp; English.

XX The present sequence is one of 3351 sequences in a library of human

CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can

CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.

XX Sequence 355 BP; 87 A; 94 C; 72 G; 102 T; 0 other;

Query Match 84.0%; Score 21; DB 22; Length 355;
Best Local Similarity 100.0%; Pred. No. 0.0052;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTCAGCCGAGATAA 21
D5 60 GGTTCAGTCAGCCGAGATAA 40
|||||

RESULT 5

ABN65341/c

ID ABN65341 standard; cDNA; 608 BP.

XX AC ABN65341;

XX 28-JUN-2002 (first entry)

DE Human cancer related polynucleotide SEQ ID NO 5308.

KW Human; cytostatic; gene expression; gene mapping; tissue profiling;
gene therapy; cancer; tumour; gene; ss.

XX Homo sapiens.

XX WO200214500-A2.

XX 21-FEB-2002.

XX 16-AUG-2001; 2001WO-US25840.

XX 16-AUG-2000; 2000US-226326P.

XX (CHIR) CHIRON CORP.

XX (HYSE-) HYSEQ INC.

PI Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F;
PI Lamson G, Scott EM, Zhang G, Kassam A, Pot D, Labat I;
XX WPI; 2002-241905/29.

XX New nucleic acid for producing a polypeptide, detecting differentially

PT expressed genes correlated with a cancerous state of a mammalian cell,
PT and inhibiting tumor growth -
XX Claim 1; SEQ ID NO 5308; 883pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (ABN27253-ABN33262)

CC with cytostatic activity. The polynucleotide is used to produce a
CC polypeptide, to detect differentially expressed genes correlated with a
CC cancerous state of a mammalian cell and to inhibit tumour growth. The
CC polynucleotide is used as a probe in mapping and tissue profiling. The
CC encoded polypeptide and antibodies to the polypeptide can also be used
CC for therapeutic and diagnostic purposes. The polynucleotide is useful for
CC gene therapy.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 608 BP; 153 A; 126 C; 114 G; 215 T; 0 other;

Query Match 84.0%; Score 21; DB 24; Length 608;
Best Local Similarity 100.0%; Pred. No. 0.0052;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGGCGGAGTAA 21
|||||
Db 501 GGTTCAGTGGCGGAGTAA 481

RESULT 6
AAK70967/c
ID AAK70967 standard; DNA; 946 BP.
XX AC AAK70967;
XX XX
DT 06-NOV-2001 (first entry)
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:25779.
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:25779.
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
XX OS
XX WO200157182-A2.
PN
XX 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 26-SEP-2000; 2000US-0234998.
PR 27-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 02-OCT-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.

PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 03-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 PI WPI; 2001-483426/52.
 XX
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis.
 XX
 XX Disclosure; SEQ ID NO 25779; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK62170 to AAK51921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK62169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 946 BP; 235 A; 243 C; 206 G; 262 T; 0 other;
 Query Match 84.0%; Score 21; DB 22; Length 946;
 Best Local Similarity 100.0%; Pred. No. 0.0052;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTTCAGTGCAGCCGAGATAA 21
 |||||||||||||||||||
 Db 899 GGTTCAGTGCAGCCGAGATAA 879
 RESULT 7
 AAH48569/c
 ID AAH48569 standard; DNA; 1242 BP.
 XX
 XX AAH48569;
 AC
 DT 20-SEP-2001 (first entry)
 XX
 DE Human fascin DNA fragment SEQ ID 19.

XX
 KW Fascin; regulatory sequence; human; dendritic cell; antiviral; tumor;
 KW antibacterial; antifungal; antiparasitic; anti-allergic; neurological;
 KW immunomodulatory; apoptotic; expression regulator; vaccine; allergen;
 KW Creutzfeld-Jakob disease; Alzheimer's disease; gene therapy;
 KW autoimmune disease; transplant rejection; ds.
 XX
 OS Homo sapiens.
 XX
 XX WO200151631-A2.
 PN
 XX
 XX 19-JUL-2001.
 PD
 XX
 XX 12-JAN-2001; 2001WO-EP00362.
 PF
 XX 13-JAN-2000; 2000DE-1001169.
 PR
 PR 02-MAR-2000; 2000DE-1010188.
 XX
 XX (RESK/) RESKE-KUNZ A.
 PA (ROSS/) ROSS X.
 PA (ROSS/) ROSS R.
 PA (BROS/) BROS M.
 XX
 XX Reske-Kunz A, Ross X, Ross R, Bros M;
 PI WPI; 2001-451859/48.
 XX
 XX New regulatory sequences from the fascin gene, useful for providing
 PT dendritic cell-specific expression of e.g. antigens, e.g. for
 PT vaccination against tumors and infections
 XX
 XX Claim 2a; Page 91; 117pp; German.
 PS
 XX
 CC This invention describes novel regulatory sequences (A) derived from
 CC human fascin that provide specific expression in dendritic cells (DC) and
 CC which have antiviral, antibacterial, antifungal, antiparasitic,
 CC anti-allergic, neurological, immunomodulatory and apoptotic activity.
 CC (A) are used to regulate expression of antigens, immunoregulators,
 CC antisense sequences etc. in DC-specific fashion. Recombinant DNA, vectors
 CC and host cells that contain (A) are useful: (i) in vaccines against
 CC viruses, bacteria, fungi, parasites, tumors, allergens and plaques in
 CC Creutzfeld-Jakob and Alzheimer's disease; and (ii) for gene therapy of
 CC tumors, allergies, infections, autoimmune diseases and transplant
 CC rejection. They can also provide specific expression of antigens and
 CC immunoregulators in DC; for isolation and identification of cell factors
 CC and cis-elements from regulatory sequences that mediate DC-specific
 CC expression; to determine the degree of maturity of DC and to block
 CC transfection factors, by providing binding sites in DC. (A) provide
 CC DC-specific expression of nucleic acid under their control, allowing a
 CC more specific regulation of the immune response and eliminating the long
 CC and laborious purification of DC (since a complete leucocyte population
 CC may be transformed), including transfection in vitro. This sequence
 CC represents a DNA fragment of human fascin which is described in the
 CC invention.
 XX
 SQ Sequence 1242 BP; 230 A; 330 C; 387 G; 295 T; 0 other;
 Query Match 84.0%; Score 21; DB 22; Length 1242;
 Best Local Similarity 100.0%; Pred. No. 0.0052;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTTCAGTGCAGCCGAGATAA 21
 |||||||||||||||||||
 Db 522 GGTTCAGTGCAGCCGAGATAA 502
 RESULT 8
 ABA92605
 ID ABA92605 standard; cDNA; 1714 BP.
 XX
 XX ABA92605;
 AC
 DT 21-MAR-2002 (first entry)
 XX

PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-213007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX (WILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 XX Schlegel R, Endege WO, Monahan JE;
 XX WPI: 2001-662795/76.
 XX
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer
 XX
 XX Claim 1; Page 5955; 11750pp; English.
 XX
 XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX
 SQ Sequence 2150 BP; 580 A; 515 C; 567 G; 485 T; 3 other;

Query Match 84.0%; Score 21; DB 23; Length 2150;
 Best Local Similarity 100.0%; Pred. No. 0.0052;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGCAGCGAGATAA 21
 ||||||||||||||||||
 Db 1777 GGTTCAGTGCAGCGAGATAA 1797

RESULT 11
 AAH48582/G
 ID AAH48582 standard; DNA; 2843 BP.
 XX
 AC AAH48582;
 XX
 DT 20-SEP-2001 (first entry)
 XX
 DE Human fascin DNA fragment SEQ ID 34.
 XX
 KW Fascin; regulatory sequence; human; dendritic cell; antiviral; tumor;
 KW antibacterial; antifungal; antiparasitic; anti-allergic; neurological;
 KW immunomodulatory; apoptotic; expression regulator; vaccine; allergen;
 KW Creutzfeld-Jakob disease; Alzheimer's disease; gene therapy;
 KW autoimmune disease; transplant rejection; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200151631-A2.
 XX
 PD 19-JUL-2001.
 XX
 PF 12-JAN-2001; 2001WO-BP00362.
 XX
 PR 13-JAN-2000; 2000DE-1001159.
 PR 02-MAR-2000; 2000DE-1010188.
 XX
 XX (RESK/) RESKE-KUNZ A.
 PA (ROSS/) ROSS X.

PA (ROSS/) ROSS R.
 PA (BROS/) BROS M.
 XX
 PI Reske-Kunz A, Ross X, Ross R, Bros M;
 XX
 DR WPI: 2001-451858/48.
 XX
 XX New regulatory sequences from the fascin gene, useful for providing
 PT dendritic cell-specific expression of e.g. antigens, e.g. for
 PT vaccination against tumors and infections
 XX
 PS Example 1; Page 103-104; 117pp; German.
 XX
 CC This invention describes novel regulatory sequences (A) derived from
 CC human fascin that provide specific expression in dendritic cells (DC) and
 CC which have antiviral, antibacterial, antifungal, antiparasitic,
 CC anti-allergic, neurological, immunomodulatory and apoptotic activity.
 CC (A) are used to regulate expression of antigens, immunoregulators,
 CC antisense sequences etc. in DC-specific fashion. Recombinant DNA, vectors
 CC and host cells that contain (A) are useful: (i) in vaccines against
 CC viruses, bacteria, fungi, parasites, tumors, allergens and plaques in
 CC Creutzfeld-Jakob and Alzheimer's disease; and (ii) for gene therapy of
 CC tumors, allergies, infections, autoimmune diseases and transplant
 CC rejection. They can also be provide specific expression of antigens and
 CC immunoregulators in DC; for isolation and identification of cell factors
 CC and cis-elements from regulatory sequences that mediate DC-specific
 CC expression; to determine the degree of maturity of DC and to block
 CC transcription factors; by providing binding sites in DC. (A) provide
 CC DC-specific expression of nucleic acid under their control, allowing a
 CC more specific regulation of the immune response and eliminating the long
 CC and laborious purification of DC (since a complete leucocyte population
 CC may be transformed), including transformation in vitro. This sequence
 CC represents a DNA fragment of human fascin which is described in the
 CC invention.
 XX
 SQ Sequence 2843 BP; 499 A; 909 C; 827 G; 608 T; 0 other;

Query Match 84.0%; Score 21; DB 22; Length 2843;
 Best Local Similarity 100.0%; Pred. No. 0.0052;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGCAGCGAGATAA 21
 ||||||||||||||||||
 Db 1607 GGTTCAGTGCAGCGAGATAA 1587

RESULT 12
 AAK84869
 ID AAK84869 standard; DNA; 5537 BP.
 XX
 AC AAK84869;
 XX
 DT 07-NOV-2001 (first entry)
 XX
 DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:39681.
 XX
 KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200157182-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01354.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226686.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 21-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 25-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.

PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-024617.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251889.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 PI
 XX WPI; 2001-483426/52.
 DR
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Disclosure; SEQ ID NO 39681; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)

CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients' own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I). Additionally, (I)
 CC the nucleic acids may be used to produce the secreted (I), by inserting
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 5537 BP; 1451 A; 1204 C; 1400 G; 1482 T; 0 other;

Query Match 84.0%; Score 21; DB 22; Length 5537;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGCAGTGCAGCCAGATAA 21
 |||||
 Db 4565 GGTGCAGTGCAGCCAGATAA 4585

RESULT 13
 AAK70965/c
 ID AAK70965 standard; DNA; 11674 BP.
 AC AAK70965;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:25777.
 XX
 KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200157182-A2.
 XX
 PD 09-AUG-2001.

XX
 PF 17-JAN-2001; 2001WO-US01354.
 XX
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216847.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 18-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 05-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 08-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.

08-NOV-2000; 2000US-0245532.
 08-NOV-2000; 2000US-0245609.
 08-NOV-2000; 2000US-0245610.
 08-NOV-2000; 2000US-0245611.
 08-NOV-2000; 2000US-0245613.
 17-NOV-2000; 2000US-0249207.
 17-NOV-2000; 2000US-0249208.
 17-NOV-2000; 2000US-0249209.
 17-NOV-2000; 2000US-0249210.
 17-NOV-2000; 2000US-0249211.
 17-NOV-2000; 2000US-0249212.
 17-NOV-2000; 2000US-0249213.
 17-NOV-2000; 2000US-0249214.
 17-NOV-2000; 2000US-0249215.
 17-NOV-2000; 2000US-0249216.
 17-NOV-2000; 2000US-0249217.
 17-NOV-2000; 2000US-0249218.
 17-NOV-2000; 2000US-0249244.
 17-NOV-2000; 2000US-0249245.
 17-NOV-2000; 2000US-0249264.
 17-NOV-2000; 2000US-0249265.
 17-NOV-2000; 2000US-0249297.
 17-NOV-2000; 2000US-0249299.
 01-DEC-2000; 2000US-0249300.
 01-DEC-2000; 2000US-0250160.
 01-DEC-2000; 2000US-0250391.
 05-DEC-2000; 2000US-0251030.
 05-DEC-2000; 2000US-0251988.
 06-DEC-2000; 2000US-0256719.
 08-DEC-2000; 2000US-0251479.
 08-DEC-2000; 2000US-0251856.
 08-DEC-2000; 2000US-0251868.
 08-DEC-2000; 2000US-0251869.
 08-DEC-2000; 2000US-0251989.
 11-DEC-2000; 2000US-0251990.
 05-JAN-2001; 2001US-0254097.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-483426/52.
 Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 useful for preventing, diagnosing and/or treating cancers and
 metastasis -
 Disclosure; SEQ ID NO 25777; 3071pp + Sequence Listing; English.
 AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic
 activity, and can be used in gene therapy and vaccine production. (I)
 proteins and polynucleotides may be used in the prevention, diagnosis and
 treatment of diseases associated with inappropriate (I) expression. For
 example, they may be used to treat disorders associated with decreased
 expression by rectifying mutations or deletions in a patient's genome
 that affect the activity of (I) by expressing inactive proteins or to
 supplement the patients own production of (I). Additionally, (I)
 polynucleotides may be used to produce the secreted (I), by inserting
 the nucleic acids into a host cell and culturing the cell to express the
 protein. (I) proteins and polynucleotides may be used to prevent,
 diagnose and treat immune/hematopoietic-related diseases, especially
 cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 to AAK87694 represent human immune/hematopoietic antigen genomic
 sequences from the present invention. AAK54942 to AAK54950 and AAK62169
 represent sequences used in the exemplification of the present invention.
 XX Sequence 11674 BP; 3026 A; 2800 C; 2616 G; 3232 T; 0 other;
 Query Match 84.0%; Score 21; DB 22; Length 11674;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGATAA 21
 |||||
 DB 899 GGTTCAGTGCAGCCGAGATAA 879
 RESULT 14
 AAH48621/c
 ID AAH48621 standard; DNA; 13055 BP.
 XX
 AC AAH48621;
 XX
 DT 20-SEP-2001 (first entry)
 XX Human fascin DNA fragment #1.
 XX
 DE 2000US-0249217.
 XX
 KW Fascin; regulatory sequence; human; dendritic cell; antiviral; tumor;
 KW antibacterial; antifungal; antiparasitic; anti-allergic; neurological;
 KW immunomodulatory; apoptotic; expression regulator; vaccine; allergen;
 KW Creutzfeld-Jakob disease; Alzheimer's disease; gene therapy;
 KW autoimmune disease; transplant rejection; ds.
 XX Homo sapiens.
 OS
 XN WO200151631-A2.
 XX
 PD 19-JUL-2001.
 XX
 XX 12-JAN-2001; 2001WO-EP00362.
 XX
 PR 13-JAN-2000; 2000DE-1001169.
 PR 02-MAR-2000; 2000DE-1010188.
 XX
 PA (RESK/) RESKE-KUNZ A.
 PA (ROSS/) ROSS X.
 PA (ROSS/) ROSS R.
 PA (BROS/) BROS M.
 XX
 PI Reske-Kunz A, Ross X, Ross R, Bros M;
 XX WPI; 2001-451858/48.
 DR
 XX
 PT New regulatory sequences from the fascin gene, useful for providing
 PT dendritic cell-specific expression of e.g. antigens, e.g. for
 PT vaccination against tumors and infections
 XX
 PS Example 1; Fig 2; 117pp; German.
 CC
 CC This invention describes novel regulatory sequences (A) derived from
 CC human fascin that provide specific expression in dendritic cells (DC) and
 CC which have antiviral, antibacterial, antifungal, antiparasitic,
 CC anti-allergic, neurological, immunomodulatory and apoptotic activity.
 CC (A) are used to regulate expression of antigens, immunoregulators,
 CC antisense sequences etc. in DC-specific fashion. Recombinant DNA, vectors
 CC and host cells that contain (A) are useful: (i) in vaccines against
 CC Creutzfeld-Jakob and Alzheimer's disease; and (ii) for gene therapy of
 CC tumors, bacteria, fungi, parasites, tumors, allergens and plaques in
 CC immunoregulators in DC; for isolation and identification of antigens and
 CC cis-elements from regulatory sequences that mediate DC-specific
 CC expression; to determine the degree of maturity of DC and to block
 CC transcription factors, by providing binding sites in DC. (A) provide
 CC DC-specific expression of nucleic acid under their control, allowing a
 CC more specific regulation of the immune response and eliminating the long
 CC and laborious purification of DC (since a complete leucocyte population
 CC may be transfected), including transformation in vitro. This sequence
 CC represents a DNA fragment of human fascin which is described in the
 CC invention.
 XX
 SQ Sequence 13055 BP; 2372 A; 3903 C; 4097 G; 2683 T; 0 other;
 Query Match 84.0%; Score 21; DB 22; Length 13055;

Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCGGAGATAA 21
|||||
Db 10885 GGTTCAGTGAGCGGAGATAA 10865

RESULT 15
AAH48620/c
ID AAH48620 standard; DNA; 16951 BP.
XX
AC AAH48620;
XX
DT 20-SEP-2001 (first entry)
XX
DE Human fascin DNA fragment SEQ ID 72.
XX
KW Fascin; regulatory sequence; human; dendritic cell; antiviral; tumor;
KW antibacterial; antifungal; antiparasitic; anti-allergic; neurological;
KW immunomodulatory; apoptotic; expression regulator; vaccine; allergen;
KW Creutzfeld-Jakob disease; Alzheimer's disease; gene therapy;
KW autoimmune disease; transplant rejection; ds.
XX
OS Homo sapiens.
XX
PN W0200151631-A2.
XX
PD 19-JUL-2001.
XX
PF 12-JAN-2001; 2001WO-EP00362.
XX
PR 13-JAN-2000; 2000DE-1001169.
PR 02-MAR-2000; 2000DE-1010188.
XX
PA (RESK/) RESKE-KUNZ A.
PA (ROSS/) ROSS X.
PA (ROSS/) ROSS R.
PA (BROS/) BROS M.
XX
PI Reske-Kunz A, Ross X, Ross R, Bros M;
XX
DR WPI; 2001-451858/48.
XX
PT New regulatory sequences from the fascin gene, useful for providing
PT dendritic cell-specific expression of e.g. antigens, e.g. for
PT vaccination against tumors and infections
XX
PS Claim 1a; Page 112-117; 117pp; German.
XX
CC This invention describes novel regulatory sequences (A) derived from
CC human fascin that provide specific expression in dendritic cells (DC) and
CC which have antiviral, antibacterial, antifungal, antiparasitic,
CC anti-allergic, neurological, immunomodulatory and apoptotic activity.
CC (A) are used to regulate expression of antigens, immunoregulators,
CC antisense sequences etc. in DC-specific fashion. Recombinant DNA, vectors
CC and host cells that contain (A) are useful: (i) in vaccines against
CC Creutzfeld-Jakob and Alzheimer's disease; and (ii) for gene therapy of
CC tumors, allergies, infections, autoimmune diseases and transplant
CC rejection. They can also be provide specific expression of antigens and
CC immunoregulators in DC; for isolation and identification of cell factors
CC and cis-elements from regulatory sequences that mediate DC-specific
CC expression; to determine the degree of maturity of DC and to block
CC DC-specific expression of nucleic acid under their control, allowing a
CC more specific regulation of the immune response and eliminating the long
CC and laborious purification of DC (since a complete leucocyte population
CC may be transformed), including transformation in vitro. This sequence
CC represents a DNA fragment of human fascin which is described in the
XX invention.
SQ Sequence 16951 BP; 3224 A; 4928 C; 5200 G; 3599 T; 0 other;

Query Match 84.0%; Score 21; DB 22; Length 16951;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCGGAGATAA 21
|||||
Db 14694 GGTTCAGTGAGCGGAGATAA 14674

RESULT 16
AAH48622/c
ID AAH48622 standard; DNA; 16951 BP.
XX
AC AAH48622;
XX
DT 20-SEP-2001 (first entry)
XX
DE Human fascin DNA fragment #2.
XX
KW Fascin; regulatory sequence; human; dendritic cell; antiviral; tumor;
KW antibacterial; antifungal; antiparasitic; anti-allergic; neurological;
KW immunomodulatory; apoptotic; expression regulator; vaccine; allergen;
KW Creutzfeld-Jakob disease; Alzheimer's disease; gene therapy;
KW autoimmune disease; transplant rejection; ds.
XX
OS Homo sapiens.
XX
PN W0200151631-A2.
XX
PD 19-JUL-2001.
XX
PF 12-JAN-2001; 2001WO-EP00362.
XX
PR 13-JAN-2000; 2000DE-1001169.
PR 02-MAR-2000; 2000DE-1010188.
XX
PA (RESK/) RESKE-KUNZ A.
PA (ROSS/) ROSS X.
PA (ROSS/) ROSS R.
PA (BROS/) BROS M.
XX
PI Reske-Kunz A, Ross X, Ross R, Bros M;
XX
DR WPI; 2001-451858/48.
XX
PT New regulatory sequences from the fascin gene, useful for providing
PT dendritic cell-specific expression of e.g. antigens, e.g. for
PT vaccination against tumors and infections
XX
PS Disclosure; Fig 9; 117pp; German.
XX
CC This invention describes novel regulatory sequences (A) derived from
CC human fascin that provide specific expression in dendritic cells (DC) and
CC which have antiviral, antibacterial, antifungal, antiparasitic,
CC anti-allergic, neurological, immunomodulatory and apoptotic activity.
CC (A) are used to regulate expression of antigens, immunoregulators,
CC antisense sequences etc. in DC-specific fashion. Recombinant DNA, vectors
CC and host cells that contain (A) are useful: (i) in vaccines against
CC Creutzfeld-Jakob and Alzheimer's disease; and (ii) for gene therapy of
CC tumors, allergies, infections, autoimmune diseases and transplant
CC rejection. They can also be provide specific expression of antigens and
CC immunoregulators in DC; for isolation and identification of cell factors
CC and cis-elements from regulatory sequences that mediate DC-specific
CC expression; to determine the degree of maturity of DC and to block
CC DC-specific expression of nucleic acid under their control, allowing a
CC more specific regulation of the immune response and eliminating the long
CC and laborious purification of DC (since a complete leucocyte population
CC may be transformed), including transformation in vitro. This sequence
CC represents a DNA fragment of human fascin which is described in the
XX invention.

```
XX SQ Sequence 16951 BP; 3224 A; 4928 C; 5200 G; 3599 T; 0 other;
Query Match      84.0%; Score 21; DB 22; Length 16951;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCGAGATAA 21
    |||||
Db 14694 GGTTCAGTGAGCGAGATAA 14674

RESULT 17
AL03002
ID AAL03002 standard; DNA; 19286 BP.
XX AC AAL03002;
XX DT 21-NOV-2001 (first entry)
XX DE Human reproductive system related antigen DNA SEQ ID NO: 5690.
XX KW Human; reproductive system related antigen; reproductive system disorder;
XX KW cancer; gene therapy; ds.
XX OS Homo sapiens.
XX PN W0200155320-A2.
XX PD 02-AUG-2001.
XX PP 17-JAN-2001; 2001WO-US01339.
XX PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225477.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234597.
PR 25-SEP-2000; 2000US-0234598.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
```

```

PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249220.
PR 17-NOV-2000; 2000US-0249221.
PR 17-NOV-2000; 2000US-0249222.
PR 17-NOV-2000; 2000US-0249223.
PR 17-NOV-2000; 2000US-0249224.
PR 17-NOV-2000; 2000US-0249225.
PR 17-NOV-2000; 2000US-0249226.
PR 17-NOV-2000; 2000US-0249227.
PR 17-NOV-2000; 2000US-0249228.
PR 17-NOV-2000; 2000US-0249229.
PR 17-NOV-2000; 2000US-0249230.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250161.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251031.
PR 05-DEC-2000; 2000US-0251032.
PR 06-DEC-2000; 2000US-0251033.
PR 06-DEC-2000; 2000US-0251034.
PR 06-DEC-2000; 2000US-0251035.
PR 08-DEC-2000; 2000US-0251036.
PR 08-DEC-2000; 2000US-0251037.
PR 08-DEC-2000; 2000US-0251038.
PR 08-DEC-2000; 2000US-0251039.
PR 08-DEC-2000; 2000US-0251040.
PR 11-DEC-2000; 2000US-0251041.
PR 11-DEC-2000; 2000US-0251042.
PR 05-JAN-2001; 2001US-0259678.
PR (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465570/50.
DR
XX
PT Isolated nucleic acid molecule encoding a reproductive system antigen
PT is used in preventing, treating or ameliorating a medical condition -
XX Disclosure; SEQ ID NO 5690; 1297pp + Sequence Listing; English.
PS
XX The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a genomic sequence encoding a
CC protein of the invention.
XX
SQ Sequence 19286 BP; 4918 A; 4223 C; 4758 G; 5387 T; 0 other;
Query Match 84.0%; Score 21; DB 22; Length 19286;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTGCGAGTGCAGCGAGATAA 21
Db 5393 GGTGCGAGTGCAGCGAGATAA 5413
RESULT 18
ABL97349
ID ABL97349 standard; DNA; 19286 BP.
XX
AC ABL97349;
XX
DT 21-JUN-2002 (first entry)
XX
DE Human testicular antigen encoding DNA fragment SEQ ID NO: 2001.
XX
KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW reproductive system disorder; urinary system disorder; gene therapy;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disease; infection; cytostatic; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200155317-A2.
XX
PD 02-AUG-2001.
XX

```

```

PR 17-JAN-2001; 2001WO-US01329.
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184564.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225256.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0234998.
PR 27-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.

```


PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PR (HUMA-) HUMAN GENOME SCI INC.
PR Rosen CA, Barash SC, Ruben SK;
PR WPI; 2001-483232/52.
PR Nucleic acids encoding 973 human testicular antigen polypeptides,
PR

PT useful for preventing, diagnosing and/or treating testicular cancer
XX Disclosure; SEQ ID NO 2001; 766pp; English.
XX
CC The present invention provides the protein and coding sequences of 973
CC human testicular antigens, and fragments of their genomic sequences. The
CC sequences can be used in the treatment of cardiovascular, urinary system,
CC reproductive system, immune, respiratory, neurological and
CC gastrointestinal disorders, infections, and particularly cancer,
CC especially testicular cancers. The present sequence is a DNA encoding a
CC protein fragment of the invention.
XX

SQ Sequence 19286 BP; 4918 A; 4223 C; 4758 G; 5387 T; 0 other;
Query Match 84.0%; Score 21; DB 23; Length 19286;
Best Local Similarity 100.0%; Pred. NO. 0.0051;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCGAGATAA 21
|||||
DB 5393 GGTTCAGTGAGCGAGATAA 5413

RESULT 19
ABAL7961
ID ABAL7961 standard; DNA; 26149 BP.
XX
AC ABAL7961;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 10292.
XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnere;
KW antiparkinsonian; antiskilling; antianaemic; antiarthritic; cancer;
KW antineumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antilucer; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX Homo sapiens.
XX
PN WO200159063-A2.
XX
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01334.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.

PR	14-AUG-2000;	2000US-0225266.
PR	14-AUG-2000;	2000US-0225270.
PR	14-AUG-2000;	2000US-0225275.
PR	14-AUG-2000;	2000US-0225277.
PR	14-AUG-2000;	2000US-0225275.
PR	14-AUG-2000;	2000US-0225258.
PR	14-AUG-2000;	2000US-0225259.
PR	18-AUG-2000;	2000US-0226279.
PR	22-AUG-2000;	2000US-0226681.
PR	22-AUG-2000;	2000US-0226682.
PR	22-AUG-2000;	2000US-0227182.
PR	23-AUG-2000;	2000US-0227009.
PR	30-SEP-2000;	2000US-0228924.
PR	01-SEP-2000;	2000US-0229287.
PR	01-SEP-2000;	2000US-0229343.
PR	01-SEP-2000;	2000US-0229344.
PR	01-SEP-2000;	2000US-0229345.
PR	05-SEP-2000;	2000US-0229509.
PR	05-SEP-2000;	2000US-0229513.
PR	06-SEP-2000;	2000US-0230438.
PR	06-SEP-2000;	2000US-0230438.
PR	08-SEP-2000;	2000US-0231242.
PR	08-SEP-2000;	2000US-0231243.
PR	08-SEP-2000;	2000US-0231244.
PR	08-SEP-2000;	2000US-0231413.
PR	08-SEP-2000;	2000US-0231414.
PR	08-SEP-2000;	2000US-0232080.
PR	08-SEP-2000;	2000US-0232081.
PR	12-SEP-2000;	2000US-0231968.
PR	14-SEP-2000;	2000US-0232397.
PR	14-SEP-2000;	2000US-0232398.
PR	14-SEP-2000;	2000US-0232399.
PR	14-SEP-2000;	2000US-0232400.
PR	14-SEP-2000;	2000US-0232401.
PR	14-SEP-2000;	2000US-0232403.
PR	14-SEP-2000;	2000US-0233064.
PR	14-SEP-2000;	2000US-0233065.
PR	21-SEP-2000;	2000US-0232223.
PR	21-SEP-2000;	2000US-0234274.
PR	25-SEP-2000;	2000US-0234997.
PR	25-SEP-2000;	2000US-0234998.
PR	26-SEP-2000;	2000US-0235484.
PR	27-SEP-2000;	2000US-0235834.
PR	27-SEP-2000;	2000US-0235836.
PR	29-SEP-2000;	2000US-0236327.
PR	29-SEP-2000;	2000US-0236367.
PR	29-SEP-2000;	2000US-0236368.
PR	29-SEP-2000;	2000US-0236369.
PR	29-SEP-2000;	2000US-0236370.
PR	02-OCT-2000;	2000US-0236802.
PR	02-OCT-2000;	2000US-0237037.
PR	02-OCT-2000;	2000US-0237038.
PR	02-OCT-2000;	2000US-0237039.
PR	13-OCT-2000;	2000US-0237040.
PR	13-OCT-2000;	2000US-0239935.
PR	20-OCT-2000;	2000US-0239937.
PR	20-OCT-2000;	2000US-0240960.
PR	20-OCT-2000;	2000US-0241785.
PR	20-OCT-2000;	2000US-0241786.
PR	20-OCT-2000;	2000US-0241787.
PR	20-OCT-2000;	2000US-0241808.
PR	20-OCT-2000;	2000US-0241809.
PR	20-OCT-2000;	2000US-0241826.
PR	01-NOV-2000;	2000US-0242221.
PR	08-NOV-2000;	2000US-0244617.
PR	08-NOV-2000;	2000US-0246474.
PR	08-NOV-2000;	2000US-0246475.
PR	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0246523.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.

08-NOV-2000; 2000US-0246527.
08-NOV-2000; 2000US-0246528.
08-NOV-2000; 2000US-0246532.
08-NOV-2000; 2000US-0246560.
08-NOV-2000; 2000US-0246609.
08-NOV-2000; 2000US-0246610.
08-NOV-2000; 2000US-0246611.
08-NOV-2000; 2000US-0246613.
17-NOV-2000; 2000US-0249207.
17-NOV-2000; 2000US-0249208.
17-NOV-2000; 2000US-0249209.
17-NOV-2000; 2000US-0249210.
17-NOV-2000; 2000US-0249211.
17-NOV-2000; 2000US-0249212.
17-NOV-2000; 2000US-0249213.
17-NOV-2000; 2000US-0249214.
17-NOV-2000; 2000US-0249215.
17-NOV-2000; 2000US-0249216.
17-NOV-2000; 2000US-0249217.
17-NOV-2000; 2000US-0249218.
17-NOV-2000; 2000US-0249244.
17-NOV-2000; 2000US-0249245.
17-NOV-2000; 2000US-0249264.
17-NOV-2000; 2000US-0249265.
17-NOV-2000; 2000US-0249297.
17-NOV-2000; 2000US-0249299.
17-NOV-2000; 2000US-0249300.
01-DEC-2000; 2000US-0250391.
01-DEC-2000; 2000US-0251160.
05-DEC-2000; 2000US-0251030.
05-DEC-2000; 2000US-0251988.
05-DEC-2000; 2000US-0256719.
06-DEC-2000; 2000US-0251479.
08-DEC-2000; 2000US-0251856.
08-DEC-2000; 2000US-0251868.
08-DEC-2000; 2000US-0251869.
08-DEC-2000; 2000US-0251989.
11-DEC-2000; 2000US-0251990.
11-DEC-2000; 2000US-0254097.
05-JAN-2001; 2001US-0259676.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;
WPI; 2001-541565/60.

Nucleic acids encoding 3224 human nervous system antigen polypeptides,
useful for preventing, diagnosing and/or treating nervous system
cancers and metastases -

Disclosure; SEQ ID NO 10292; 1701pp + Sequence Listing; English.

The invention relates to novel genes (ABAI1004-ABA21534) and proteins
(ABBI14678-ABBI14601) useful for preventing, treating or ameliorating
medical conditions e.g. by protein or gene therapy. The genes are
isolated from a range of human tissues disclosed in the specification.
The nucleic acids, proteins, antibodies and (ant)agonists are useful
in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
and ovarian cancer and other cancers of the adrenal gland, bone, bone
marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
(b) immune disorders e.g. Addison's disease, allergies, autoimmune
haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
disease, multiple sclerosis, rheumatoid arthritis and ulcerative
colitis; (c) cardiovascular disorders such as myocardial ischaemias;
(d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
and parasitic infections.

Note: The sequence data for this patent did not form part of the
printed specification, but was obtained in electronic format directly
from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 28149 BP; 7151 A; 6467 C; 6967 G; 7564 T; 0 other:

Query Match 84.0%; Score 21; DB 22; Length 28149;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCGAGATAA 21
 |||||
 Db 2876 GGTTCAGTGCAGCGAGATAA 2896

RESULT 20
 AAK82759/c
 ID AAK82759 standard; DNA; 30360 BP.
 AC AAK82759;

XX 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:37571.
 XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis; ds.
 XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01354.

PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216680.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 05-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 06-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0233297.
 PR 14-SEP-2000; 2000US-0233398.
 PR 14-SEP-2000; 2000US-0233399.
 PR 14-SEP-2000; 2000US-0234400.
 PR 14-SEP-2000; 2000US-0234401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 21-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 25-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 26-SEP-2000; 2000US-0234998.
 PR 27-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 29-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 20-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 08-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.

PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0254097.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI; 2001-483426/52.
 XX
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis.
 XX
 PS Disclosure; SEQ ID NO 37571; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK62169
 CC represent sequences used in the exemplification of the present invention.
 XX
 XX Sequence 30360 BP; 8943 A; 5665 C; 6045 G; 9707 T; 0 other;
 SQ
 Query Match 84.0%; Score 21; DB 22; Length 30360;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTTCAGTGGCCGAGATAA 21
 Db 14433 GGTTCAGTGGCCGAGATAA 14413
 RESULT 21
 ID ABA17963
 XX
 XX ID ABA17963 standard; DNA; 32192 BP.
 AC
 AC ABA17963;
 XX
 XX 23-JAN-2002 (first entry)
 DT
 XX Human nervous system related polynucleotide SEQ ID NO 10294.
 DE
 XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
 KW

KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
 KW antiparkinsonian; antisickling; antianaemic; antiarthritis; cancer;
 KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
 KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
 KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
 XX neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
 OS Homo sapiens.
 XX
 XX WO200159063-A2.
 PN
 XX
 PD 16-AUG-2001.
 XX
 XX 17-JAN-2001; 2001WO-US01334.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.

PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234224.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0234999.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-541565/60.
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system
PT cancers and metastases -
XX Disclosure; SEQ ID NO 10294; 1701pp + Sequence Listing; English.
XX The invention relates to novel genes (AB11004-ABA21534) and proteins
CC (AB114678-AB118001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 32192 BP; 7965 A; 7802 C; 8055 G; 8370 T; 0 other;
Query Match 84.0%; Score 21; DB 22; Length 32192;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTTCAGCTGAGCCGAGATAA 21
|||||
Db 2876 GGTTCAGCTGAGCCGAGATAA 2896
RESULT 22
AAS98633
ID AAS98633 standard; DNA; 38258 BP.
XX
AC AAS98633;
XX
DT 26-MAR-2002 (first entry)
XX
DE DNA encoding Colony stimulating factor 1 receptor (CSF1R).
XX Colony stimulating factor 1 receptor; CSF1R; polymorphic variant;
KW cytosolic; gene therapy; malignant histiocytosis; isogene;
KW myeloid malignancy; inflammatory disorder; transgenic animal;
XX haplotype; genotype; human; chromosome 5q33-35; ds.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FT variation 867
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT variation 875
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism"
FT variation 1022

```

FT      /*tag= c
FT      /standard_name= "Single nucleotide polymorphism"
FT      1076
FT      /tag= d
FT      /standard_name= "Single nucleotide polymorphism"
FT      1174
FT      /tag= e
FT      /standard_name= "Single nucleotide polymorphism"
FT      1302
FT      /tag= f
FT      /standard_name= "Single nucleotide polymorphism"
FT      1591
FT      /tag= g
FT      /standard_name= "Single nucleotide polymorphism"
FT      1617
FT      /tag= h
FT      /standard_name= "Single nucleotide polymorphism"
FT      1810
FT      /tag= i
FT      /standard_name= "Single nucleotide polymorphism"
FT      1916..34258
FT      /tag= j
FT      /standard_name= "Single nucleotide polymorphism"
FT      /product= "CSFIR"
FT      /note= "Colony stimulating factor 1 receptor. Coding
FT      sequence is specifically claimed in claim 24"
FT      1916..1964
FT      /tag= k
FT      /number= 1
FT      1965..7315
FT      /tag= l
FT      /number= 1
FT      2015
FT      /tag= m
FT      /standard_name= "Single nucleotide polymorphism"
FT      2043
FT      /tag= n
FT      /standard_name= "Single nucleotide polymorphism"
FT      7314..7571
FT      /tag= o
FT      /number= 2
FT      7348
FT      /tag= p
FT      /standard_name= "Single nucleotide polymorphism"
FT      7532
FT      /tag= q
FT      /standard_name= "Single nucleotide polymorphism"
FT      7546
FT      /tag= r
FT      /standard_name= "Single nucleotide polymorphism"
FT      7558
FT      /tag= s
FT      /standard_name= "Single nucleotide polymorphism"
FT      7572..8001
FT      /tag= t
FT      /number= 2
FT      8002..8286
FT      /tag= u
FT      /number= 3
FT      8062
FT      /tag= v
FT      /standard_name= "Single nucleotide polymorphism"
FT      8151
FT      /tag= w
FT      /standard_name= "Single nucleotide polymorphism"
FT      8287..10089
FT      /tag= x
FT      /number= 3
FT      8327
FT      /tag= y
FT      /standard_name= "Single nucleotide polymorphism"
FT      8340
FT      /tag= z
FT      /standard_name= "Single nucleotide polymorphism"

```

```

FT      variation
FT      /*tag= aa
FT      /standard_name= "Single nucleotide polymorphism"
FT      10082
FT      /tag= ab
FT      /standard_name= "Single nucleotide polymorphism"
FT      10090..10226
FT      /tag= ac
FT      /number= 4
FT      10218
FT      /tag= ad
FT      /standard_name= "Single nucleotide polymorphism"
FT      10223
FT      /tag= ae
FT      /standard_name= "Single nucleotide polymorphism"
FT      10227..10902
FT      /tag= af
FT      /number= 4
FT      10279
FT      /tag= ag
FT      /standard_name= "Single nucleotide polymorphism"
FT      11008
FT      /tag= aj
FT      /standard_name= "Single nucleotide polymorphism"
FT      11063..14842
FT      /tag= ak
FT      /number= 5
FT      11090
FT      /tag= al
FT      /standard_name= "Single nucleotide polymorphism"
FT      11129
FT      /tag= am
FT      /standard_name= "Single nucleotide polymorphism"
FT      11130
FT      /tag= an
FT      /standard_name= "Single nucleotide polymorphism"
FT      11207
FT      /tag= ao
FT      /standard_name= "Single nucleotide polymorphism"
FT      14843..15035
FT      /tag= ap
FT      /number= 6
FT      15036..17765
FT      /tag= aq
FT      /number= 6
FT      15068
FT      /tag= ar
FT      /standard_name= "Single nucleotide polymorphism"
FT      15106
FT      /tag= as
FT      /standard_name= "Single nucleotide polymorphism"
FT      17642
FT      /tag= at
FT      /standard_name= "Single nucleotide polymorphism"
FT      17758
FT      /tag= au
FT      /standard_name= "Single nucleotide polymorphism"
FT      17766..17881
FT      /tag= av
FT      /number= 7
FT      17768
FT      /tag= aw
FT      /standard_name= "Single nucleotide polymorphism"
FT      17882..18034
FT      /tag= ax
FT      /number= 7
FT      18035..18155
FT      exon

```

```

FT      /*tag= ay
FT      /number= 8
FT      intron      18156..18273
FT      /*tag= az
FT      /number= 8
FT      exon      18274..18464
FT      /*tag= ba
FT      /number= 9
FT      intron      18465..20006
FT      /*tag= bb
FT      exon      20007..20122
FT      /*tag= bc
FT      /number= 10
FT      intron      20123..26477
FT      /*tag= bd
FT      /number= 10
FT      variation      20125
FT      /*tag= be
FT      /standard_name= "Single nucleotide polymorphism"
FT      20129
FT      /*tag= bf
FT      /standard_name= "Single nucleotide polymorphism"
FT      20134
FT      /*tag= bg
FT      /standard_name= "Single nucleotide polymorphism"
FT      2013
FT      /*tag= bh
FT      /standard_name= "Single nucleotide polymorphism"
FT      20272
FT      /*tag= bi

```

Query Match 84.0%; Score 21; DB 24; Length 38258;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 GGTTCAGTGAGCGAGATAA 21
      |||||
Db      37789 GGTTCAGTGAGCGAGATAA 37809

```

```

RESULT 23
AAZ23903
ID      AAZ23903 standard; DNA; 49999 BP.
XX      AC
XX      AAZ23903;
XX      25-JAN-2000 (first entry)
XX      Human LOBO homologue genomic DNA fragment 5.
XX      LOBO; long bones; bone development; bone extension; skull; osteopathic;
XX      diagnostic; pharmaceutical; gene therapy; transgenic animal; disease;
XX      spondyloepiphyseal dysplasia; achondroplasia; human; ds.
XX      Homo sapiens.
XX      W09950284-A2.
XX      07-OCT-1999.
XX      26-MAR-1999; 99WO-EP02055.
XX      27-MAR-1998; 98DE-1013799.
XX      (ROSE/) ROSENTHAL A.
XX      Rosenthal A, Rump A, Hess J, Aigner T, Wirth T;
XX      WPI; 1999-601320/51.
XX      Nucleic acids encoding proteins which influence bone development,
XX      useful for treating and studying bone disorders -

```

```

XX      PS      Example 3; Page 328-356; 391pp; German.
XX      CC      This invention describes novel nucleic acids (I; designated LOBO (long
XX      CC      bones)) encoding proteins influencing bone development in mammals. The
XX      CC      proteins of the invention reduce and/or inactivate bone extension (i.e.
XX      CC      development), with exception of the skull and have osteopathic activity.
XX      CC      The nucleic acid molecules, proteins and antibodies can be used in
XX      CC      diagnostic or pharmaceutical compounds e.g. for gene therapy. The methods
XX      CC      and nucleic acid molecules, etc. are useful for production of transgenic
XX      CC      animals, especially a transgenic mouse for the study of diseases
XX      CC      associated with bone development, e.g. spondyloepiphyseal dysplasia and
XX      CC      achondroplasia. This sequence encodes a human LOBO protein described
XX      CC      in the method of the invention.
SQ      Sequence 49999 BP; 10983 A; 13723 C; 13439 G; 11854 T; 0 other;

```

Query Match 84.0%; Score 21; DB 20; Length 49999;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 GGTTCAGTGAGCGAGATAA 21
      |||||
Db      30972 GGTTCAGTGAGCGAGATAA 30992

```

```

RESULT 24
AAC15906/c
ID      AAC15906 standard; cDNA; 95 BP.
XX      AC
XX      AAC15906;
XX      06-OCT-2000 (first entry)
XX      Human secreted protein 5' EST, SEQ ID NO: 19981.
XX      Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX      gene therapy; chromosome mapping; ss.
XX      Homo sapiens.
XX      EP1033401-A2.
XX      06-SEP-2000.
XX      21-FEB-2000; 2000EP-0200610.
XX      26-FEB-1999; 99US-0122487.
XX      (GEST ) GENSET.
XX      Dumas Milne Edwards J, Duclert A, Giordano J;
XX      WPI; 2000-500381/45.
XX      New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX      obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
XX      diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX      Claim 1; SEQ ID 19981; 71pp + CD-ROM; English.
XX      The present sequence is one of a large number of 5' ESTs derived from
XX      mRNAs encoding secreted proteins. No ORF has yet been conclusively
XX      identified within the present sequence. The 5' ESTs were prepared from
XX      total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
XX      sequences usually correspond mainly to the 3' untranslated region (UTR)
XX      of the mRNA because they are often obtained from oligo-dT primed cDNA
XX      libraries. Such ESTs are not well suited for isolating cDNA sequences
XX      derived from the 5' ends of mRNAs and even in those cases where longer
XX      cDNA sequences have been obtained, the full 5' UTR is rarely included.
XX      5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
XX      used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
XX      in diagnostic, forensic, gene therapy and chromosome mapping procedures.

```

CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.

SQ Sequence 95 BP; 11 A; 29 C; 21 G; 34 T; 0 other;

Query Match 80.0%; Score 20; DB 21; Length 95;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCGAGTGCAGGAGATA 20
Db 90 GGTTCGAGTGCAGGAGATA 71

RESULT 25

AAK81714/c
ID AAK81714 standard; DNA; 171 BP.

XX AAK81714;

XX 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:36526.

KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytosolic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX W0200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 07-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.

XX 11-JUL-2000; 2000US-0217496.

XX 14-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220963.

XX 26-JUL-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224518.

XX 14-AUG-2000; 2000US-0224519.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225214.

XX 14-AUG-2000; 2000US-0225266.

XX 14-AUG-2000; 2000US-0225267.

XX 14-AUG-2000; 2000US-0225268.

XX 14-AUG-2000; 2000US-0225270.

XX 14-AUG-2000; 2000US-0225270.

XX 14-AUG-2000; 2000US-0225447.

XX 14-AUG-2000; 2000US-0225757.

XX 14-AUG-2000; 2000US-0225758.

XX 18-AUG-2000; 2000US-0225759.

XX 22-AUG-2000; 2000US-0226279.

XX 22-AUG-2000; 2000US-0226681.

XX 22-AUG-2000; 2000US-0226686.

XX 23-AUG-2000; 2000US-0227182.

XX 30-AUG-2000; 2000US-0227009.

XX 01-SEP-2000; 2000US-0228924.

XX 01-SEP-2000; 2000US-0229287.

XX 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0241826.
PR 08-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0244674.
PR 08-NOV-2000; 2000US-0244675.
PR 08-NOV-2000; 2000US-0244676.
PR 08-NOV-2000; 2000US-0244677.
PR 08-NOV-2000; 2000US-0244678.
PR 08-NOV-2000; 2000US-0244652.
PR 08-NOV-2000; 2000US-0244653.
PR 08-NOV-2000; 2000US-0244654.
PR 08-NOV-2000; 2000US-0244655.
PR 08-NOV-2000; 2000US-0244656.
PR 08-NOV-2000; 2000US-0244657.
PR 08-NOV-2000; 2000US-0244658.
PR 08-NOV-2000; 2000US-0244659.
PR 08-NOV-2000; 2000US-0244660.
PR 08-NOV-2000; 2000US-0244661.
PR 08-NOV-2000; 2000US-0244663.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.

PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249267.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250169.
 PR 05-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-483426/52.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Disclosure; SEQ ID NO 36526; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC the nucleic acids may be used to produce the secreted (I), by inserting
 CC protein, (I) proteins and polynucleotides may be used to prevent,
 CC cancers and treat immune/hematopoietic-related diseases, especially
 CC to AAK67694 represent human immune/hematopoietic-derived cells. AAK64703
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 171 BP; 28 A; 52 C; 39 G; 52 T; 0 other;

 Query Match 80.0%; Score 20; DB 22; Length 171;
 Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 2 GTTCAGCTGAGCCGAGATAA 21
 Db |||||
 79 GTTCAGCTGAGCCGAGATAA 60

 RESULT 26
 AAK83823
 ID AAK83823 standard; DNA; 172 BP.
 XX
 AC AAK83823;
 XX

DT 07-NOV-2001 (first entry)
 XX Human immune/hematopoietic antigen genomic sequence SEQ ID NO:38635.
 DE Human; immune; haematopoietic; immune/hematopoietic antigen; cancer;
 XX cytostatic; gene therapy; vaccine; metastasis; ds.
 KW Homo sapiens.
 KW
 OS WO200157182-A2.
 XX
 PN 09-AUG-2001.
 XX
 PD 17-JAN-2001; 2001WO-US01354.
 XX
 PF 31-JAN-2000; 2000US-0179065.
 XX 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0215880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 23-AUG-2000; 2000US-0227182.
 PR 30-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.

PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234224.
 PR 25-SEP-2000; 2000US-023497.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 29-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 02-OCT-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 20-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 08-NOV-2000; 2000US-024617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249267.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 05-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251368.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.

PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251889.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-483426/52.
 XX Nucleic acids encoding human immune/haematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis.
 XX Disclosure; SEQ ID NO 38635; 3071pp + Sequence Listing; English.
 XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK34942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX Sequence 172 BP; 56 A; 38 C; 49 G; 29 T; 0 other;
 SQ
 Query Match 80.0%; Score 20; DB 22; Length 172;
 Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTTCAGTGCAGCCGAGATA 20
 Db 81 GGTTCAGTGCAGCCGAGATA 100
 RESULT 27
 AAC11950/c
 ID AAC11950 standard; cDNA; 180 BP.
 XX AAC11950;
 AC AAC11950;
 XX 06-OCT-2000 (first entry)
 DT Human secreted protein 5' EST, SEQ ID NO: 16025.
 DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 XX gene therapy; chromosome mapping; ss.
 OS Homo sapiens.
 XX EP1033401-A2.
 XX 06-SEP-2000.
 PD 21-FEB-2000; 2000EP-0200610.
 XX 26-FEB-1999; 99US-0122487.
 PR (GEST) GENSET.
 PA Dumas Milne Edwards J, Duclert A, Giordano J;
 PI

XX WPI; 2000-500381/45.
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX Claim 1; SEQ ID 16025; 71pp + CD-ROM; English.
 XX The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.
 XX Sequence 180 BP; 42 A; 59 C; 52 G; 27 T; 0 other;
 SQ
 Query Match 80.0%; Score 20; DB 21; Length 180;
 Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTTCAGTGCAGCGAGATA 20
 DB 83 GGTTCAGTGCAGCGAGATA 64
 RESULT 28
 AAC04572/C
 ID AAC04572 standard; cDNA; 193 BP.
 XX AAC04572;
 AC
 XX 06-OCT-2000 (first entry)
 DT
 DE Human secreted protein 5' EST, SEQ ID NO: 8647.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.
 XX Homo sapiens.
 OS
 XX EP1033401-A2.
 FN
 XX 06-SEP-2000.
 PD
 XX 21-FEB-2000; 2000EP-0200610.
 PF
 XX 26-FEB-1999; 99US-0122487.
 PR
 XX (GIST) GENSET.
 PA
 XX Dumas Milne Edwards J, Duclert A, Giordano J;
 PI WPI; 2000-500381/45.
 XX
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX Claim 1; SEQ ID 8647; 71pp + CD-ROM; English.
 XX The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from

CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.
 XX Sequence 193 BP; 37 A; 41 C; 50 G; 65 T; 0 other;
 SQ
 Query Match 80.0%; Score 20; DB 21; Length 193;
 Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTTCAGTGCAGCGAGATA 20
 DB 155 GGTTCAGTGCAGCGAGATA 136
 RESULT 29
 AAI68895/C
 ID AAI68895 standard; DNA; 205 BP.
 XX AAI68895;
 AC
 XX 29-JAN-2002 (first entry)
 DT
 DE Activated T-cell derived DNA fragment #31.
 XX
 KW Activated T-cell; immunosuppressive; immunostimulant; antiinflammatory;
 KW cytostatic; gene therapy; vaccine; allergen; transplant rejection;
 KW guest versus host disease; malignant disease; ds.
 XX Homo sapiens.
 OS
 XX DE10021834-A1.
 FN
 XX 15-NOV-2001.
 PD
 XX 06-MAY-2000; 2000DE-1021834.
 PF
 XX 06-MAY-2000; 2000DE-1021834.
 PR
 XX (LYNX-) LYNX THERAPEUTICS GMBH.
 PA
 XX Kramer MD, Winter H, Reinartz J;
 PI WPI; 2002-027320/04.
 DR
 XX New mRNA indicative of T cell activation and functional status, useful
 PT for diagnosis and therapy e.g. of autoimmunity or transplant rejection
 PT
 XX Claim 1; Page 17; 94pp; German.
 PS
 XX This sequence represents a novel messenger RNA, (mRNA), (I), for use as
 CC indicator of the activation and functional status of T cells, that have
 CC increased or reduced expression, and are present at higher or lower
 CC concentration, in activated T cells, relative to normal or resting cells,
 CC where (I) hybridizes to any of 334 sequences, reproduced, or their
 CC derivatives, complements or fragments. The products of the invention
 CC have immunosuppressive, immunostimulant, antiinflammatory and cytostatic
 CC activity and can be used for gene therapy. The polynucleotides of the
 CC invention are used: (i) as reagent for detecting activation/functional
 CC status of T cells, for diagnosis, therapy, modulation or control of the
 CC status, in cases of (auto)immunity (against microorganisms, vaccines or
 CC allergens); transplant rejection; immunologically-related inflammation;
 CC immunosuppression; immune deficiency; guest versus host disease, and
 CC malignant diseases of the immune system; (ii) for identifying agents,

CC potential pharmaceuticals, that bind to (II) or derived polypeptides
 CC (III); (iii) to prepare kits for measuring gene expression profiles in
 CC isolated immune, especially T, cells; (iv) to raise antibodies (Ab)
 CC directed against (III); and (v) to prepare binding molecules (IV)
 CC specific for (II). Ab and (IV) are also useful for detecting and
 CC modulating the activation and functional status of T cells.
 CC AAI6865-AA169198 represent the activated T-cell derived polynucleotide
 CC fragments described in the method of the invention.

XX SQ Sequence 205 BP; 37 A; 54 C; 55 G; 59 T; 0 other;

Query Match 80.0%; Score 20; DB 24; Length 205;
 Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGATA 20
 DB 64 GGTTCAGTGCAGCCGAGATA 45
 |||||

RESULT 30

AAC32484/c
 ID AAC32484 standard; cDNA; 236 BP.

XX AAC32484;

DT 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 36559.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

PF 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI: 2000-500381/45.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX Claim 1; SEQ ID 36559; 71pp + CD-ROM; English.

CC The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.

SQ Sequence 236 BP; 58 A; 47 C; 55 G; 76 T; 0 other;

Query Match 80.0%; Score 20; DB 21; Length 236;

Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGATA 20
 DB 96 GGTTCAGTGCAGCCGAGATA 77
 |||||

RESULT 31

AAC04849/c

ID AAC04849 standard; cDNA; 304 BP.

XX AAC04849;

DT 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 8924.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

PF 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI: 2000-500381/45.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX Claim 1; SEQ ID 8924; 71pp + CD-ROM; English.

CC The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.

SQ Sequence 304 BP; 45 A; 80 C; 77 G; 90 T; 12 other;

Query Match 80.0%; Score 20; DB 21; Length 304;
 Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGATA 20
 DB 87 GGTTCAGTGCAGCCGAGATA 68
 |||||

RESULT 32

AAH68890

ID AAH68890 standard; cDNA; 320 BP.

XX

```
AC AAH68890;
XX
DT 19-SEP-2001 (first entry)
XX
DE Human cervical cancer marker nucleic acid 164.
XX
KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200142467-A2.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-US33312.
XX
PR 08-DEC-1999; 99US-0169681.
XX
PR 21-DEC-1999; 99US-0171350.
XX
PR 14-MAR-2000; 2000US-0189315.
XX
PR 12-MAY-2000; 2000US-0203791.
XX
PR 09-JUN-2000; 2000US-0210600.
XX
PR 21-JUL-2000; 2000US-0220114.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Deeds J, Berger A, Zhao X;
XX
XX WPI; 2001-375006/39.
XX
XX New isolated nucleic acid for diagnosing and treating cervical cancer
XX and for assessing and detecting compounds for treating the cancer -
XX
XX Claim 1; Page 131; 1051pp; English.
XX
XX The invention relates to novel genes (AAH68727-AAH73383) associated with
XX cervical cancer with cytostatic activity. The nucleic acids and encoded
XX polypeptides are useful: to assess if a patient is afflicted with
XX cervical cancer or has a pre-malignant condition; to monitor the
XX progression of cervical cancer or a premalignant condition in a patient;
XX and to select and/or assess the efficacy of a compound or therapy for
XX inhibiting cervical cancer in a patient. The nucleic acids may also be
XX useful for gene therapy.
XX
XX Sequence 320 BP; 96 A; 67 C; 86 G; 71 T; 0 other;
XX
XX Query Match 80.0%; Score 20; DB 22; Length 320;
XX Best Local Similarity 100.0%; Pred. No. 0.02;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 GGTTCAGTGCAGCCGAGATA 20
XX |||||||||||||||||||
XX Db 226 GGTTCAGTGCAGCCGAGATA 245
XX
XX RESULT 33
XX AAH70549
XX ID AAH70549 standard; cDNA; 329 BP.
XX
XX AC AAH70549;
XX
XX 19-SEP-2001 (first entry)
XX
XX Human cervical cancer marker nucleic acid 1823.
XX
XX Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
XX Homo sapiens.
XX
XX WO200142467-A2.
XX
XX 14-JUN-2001.
XX
XX 08-DEC-2000; 2000WO-US33312.
XX
XX PF
```

```
XX
PR 08-DEC-1999; 99US-0169681.
XX
PR 21-DEC-1999; 99US-0171350.
XX
PR 14-MAR-2000; 2000US-0189315.
XX
PR 12-MAY-2000; 2000US-0203791.
XX
PR 09-JUN-2000; 2000US-0210600.
XX
PR 21-JUL-2000; 2000US-0220114.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Deeds J, Berger A, Zhao X;
XX
XX WPI; 2001-375006/39.
XX
XX New isolated nucleic acid for diagnosing and treating cervical cancer
XX and for assessing and detecting compounds for treating the cancer -
XX
XX Claim 1; Page 394; 1051pp; English.
XX
XX The invention relates to novel genes (AAH68727-AAH73383) associated with
XX cervical cancer with cytostatic activity. The nucleic acids and encoded
XX polypeptides are useful: to assess if a patient is afflicted with
XX cervical cancer or has a pre-malignant condition; to monitor the
XX progression of cervical cancer or a premalignant condition in a patient;
XX and to select and/or assess the efficacy of a compound or therapy for
XX inhibiting cervical cancer in a patient. The nucleic acids may also be
XX useful for gene therapy.
XX
XX Sequence 329 BP; 99 A; 68 C; 88 G; 74 T; 0 other;
XX
XX Query Match 80.0%; Score 20; DB 22; Length 329;
XX Best Local Similarity 100.0%; Pred. No. 0.02;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 GGTTCAGTGCAGCCGAGATA 20
XX |||||||||||||||||||
XX Db 235 GGTTCAGTGCAGCCGAGATA 254
XX
XX RESULT 34
XX AAH7614
XX ID AAH7614 standard; cDNA; 335 BP.
XX
XX AC AAH7614;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 7674.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation; ss.
XX
XX Homo sapiens.
XX
XX WO200164835-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US04927.
XX
XX 28-FEB-2000; 2000US-0515126.
XX
XX 18-MAY-2000; 2000US-0577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-514838/56.
XX
XX P-PSDB; AAO07683.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing
XX PT
```

PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -
 XX
 PS Claim 1; SEQ ID NO 7674; 1399pp + Sequence Listing; English.
 XX
 CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 335 BP; 117 A; 54 C; 84 G; 79 T; 1 other;
 SQ

Query Match 80.0%; Score 20; DB 22; Length 335;
 Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGTTCAGTGCAGCCGAGATA 20
 |||||
 DB 11 GGTTCAGTGCAGCCGAGATA 30

RESULT 35

ABL79179/C
 ID ABL79179 standard; cDNA; 384 BP.

XX
 AC ABL79179;

DT 17-MAY-2002 (first entry)

DE Human ovarian cancer related cDNA clone SEQ ID NO:2157.
 DE
 KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.

XX Homo sapiens.
 OS
 XX WO200192581-A2.

PN
 XX
 XX PD 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US17756.

XX 26-MAY-2000; 2000US-207484P.

XX (CORI-) CORIXA CORP.

XX
 XX PI Algate PA, Harlocker SL, Jones R;

XX WPI; 2002-122075/16.

XX
 XX PT Composition for therapy and diagnosis of ovarian cancer comprising

PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
 PT polypeptide, antibody specific to polypeptide or T cell expressing
 PT polypeptide -

XX
 XX PS Claim 1; SEQ ID 2157; 489pp; English.

XX The present invention describes a composition (I) comprising: carriers
 CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
 CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence
 CC (SI) from the 10912 nucleotide sequences as given in ABL77023 to
 CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
 CC population of (II), or antigen presenting cells that express (II).
 CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to

CC (SI) can be used for detecting ovarian cancer in a patient's biological
 CC sample preferably serum or ovarian tissue. The method comprises
 CC contacting a biological sample from a patient with (IV), detecting the
 CC amount of polynucleotide hybridising to (IV) and comparing the amount to
 CC a predetermined cutoff value and thereby detecting ovarian cancer in the
 CC patient, where the amount of polynucleotide hybridising to (IV) is
 CC detected preferably by polymerase chain reaction (PCR). (I) comprising
 CC (III) and/or (II) is useful for stimulating and/or expanding T cells
 CC specific for an ovarian tumour protein comprising contacting T cells
 CC with (III) or (II). (III) is useful in design and preparation of
 CC ribozyme molecules for inhibiting expression of the tumour polypeptides
 CC and proteins in tumour cells; and to isolate a full length gene from a
 CC suitable library e.g., a tumour cDNA library using well known
 CC techniques.

XX SQ Sequence 384 BP; 90 A; 102 C; 89 G; 100 T; 3 other;
 SQ

Query Match 80.0%; Score 20; DB 24; Length 384;
 Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGTTCAGTGCAGCCGAGATA 20
 |||||
 DB 70 GGTTCAGTGCAGCCGAGATA 51

RESULT 36

AAI82847
 ID AAI82847 standard; cDNA; 388 BP.

XX
 AC AAI82847;

XX 06-NOV-2001 (first entry)

XX Human polynucleotide SEQ ID NO 2907.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation; ss.

XX Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US04927.

XX 28-FEB-2000; 2000US-0515126.

XX 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

XX P-PSDB; AAC02916.

XX Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -

XX Claim 1; SEQ ID NO 2907; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 388 BP; 94 A; 52 C; 63 G; 179 T; 0 other;

Query Match 80.0%; Score 20; DB 22; Length 388;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCGAGATA 20
|||||
Db 55 GGTTCAGTGCAGCGAGATA 74

RESULT 37

ABV16660
ID ABV16660 standard; cDNA; 425 BP.

XX AC ABV16660;

XX 13-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 16651.

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US05171.

XX PR 17-FEB-2000; 2000US-183319P.

XX PR 16-MAR-2000; 2000US-189862P.

XX PR 25-MAY-2000; 2000US-207454P.

XX PR 09-JUN-2000; 2000US-211314P.

XX PR 18-JUL-2000; 2000US-219007P.

XX PR 13-DEC-2000; 2000US-255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX PS WPI; 2001-662795/76.

XX PT Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.

XX PS Claim 1; Page 2775; 11/750pp; English.

XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for:

XX CC (a) assessing whether a patient is afflicted with prostate cancer;

XX CC (b) monitoring the progression of prostate cancer in a patient;

XX CC (c) assessing the efficacy of a test compound to inhibit prostate
XX cancer in a patient;

XX CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX in a patient;

XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;

XX CC (f) assessing the prostate cell carcinogenic potential of a compound;

XX CC (g) determining whether prostate cancer has metastasized in a patient;

XX CC (h) assessing the aggressiveness or incidence of prostate cancer in a
XX patient;

XX CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX SQ

Sequence 425 BP; 150 A; 96 C; 92 G; 87 T; 0 other;

Query Match 80.0%; Score 20; DB 23; Length 425;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCGAGATA 20
|||||

Db 198 GGTTCAGTGCAGCGAGATA 217
|||||

RESULT 38

AAI83705

ID AAI83705 standard; cDNA; 428 BP.

XX AC AAI83705;

XX 06-NOV-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 3765.

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.

XX OS Homo sapiens.

XX PN WO200164835-A2.

XX PD 07-SEP-2001.

XX PF 26-FEB-2001; 2001WO-US04927.

XX PR 28-FEB-2000; 2000US-0515126.

XX PR 18-MAY-2000; 2000US-0577409.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YI, Liu C, Drmanac RT;

XX DR WPI; 2001-514838/56.

XX DR P-PSDB; AAO03774.

XX PT Isolated nucleic acids and polypeptides, useful for preventing
XX diagnosing and treating e.g. leukaemia, inflammation and immune
XX disorders -

XX PS Claim 1; SEQ ID NO 3765; 1399pp + Sequence Listing; English.

XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX the encoded proteins (AAO00010-AAO1910) that exhibit activity relating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation.

XX CC Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 428 BP; 108 A; 102 C; 130 G; 86 T; 2 other;

Query Match 80.0%; Score 20; DB 22; Length 428;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCGAGATA 20

Db 34 GGTTCAGTGGCCGAGATA 53
|||||
RESULT 39
ABV49562
ID ABV49562 standard; cDNA; 436 BP.
XX
XX
AC ABV49562;
DT 17-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 49553.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW Pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US05171.
XX
XX 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
PI
XX
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer -
XX
XX Claim 1; Page 9674; 11750pp; English.
CC The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 436 BP; 138 A; 102 C; 122 G; 74 T; 0 other;
Query Match 80.0%; Score 20; DB 23; Length 436;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTTCAGTGGCCGAGATA 20
|||||
Db 232 GGTTCAGTGGCCGAGATA 251
RESULT 40
ABQ89507/c

ID ABQ89507 standard; cDNA; 436 BP.
XX
XX AC ABQ89507;
XX
XX 27-SEP-2002 (first entry)
XX
XX Human prostate expressed polynucleotide SEQ ID NO 763.
XX
XX Human; prostate; cytostatic; tumour; cancer; vaccine; gene therapy;
KW gene; ss.
XX
XX Homo sapiens.
XX
XX WO200255700-A2.
XX
XX 18-JUL-2002.
XX
XX 07-DEC-2001; 2001WO-US47349.
XX
XX 07-DEC-2000; 2000US-254648P.
PR 13-MAR-2001; 2001US-275688P.
XX
XX (CHIR) CHIRON CORP.
XX (HYSE-) HYSEQ INC.
XX
XX Escobedo J, Garcia PD, Kassam A, Lamson G, Drmanac R;
PI Ctkvenjakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones WL, Stache-Crain B, Scott EM,
XX WPI; 2002-557824/59.
XX
XX New genes and gene products isolated from human prostate, useful for treating or diagnosing tumor or cancer (e.g. prostate cancer or breast cancer), or as vaccines for treating or preventing these diseases -
XX
XX Claim 1; SEQ ID NO 763; 186pp + Sequence Listing; English.
XX
XX The invention relates to an isolated polynucleotide comprising any of 1477 sequences or its fragment, degenerate variant, antisense or complement. The polynucleotides and gene products are useful for treating or diagnosing tumour or cancer (e.g. prostate cancer, breast cancer, lung cancer or medullary carcinoma) in a subject (e.g. cattle, dogs, cats, rabbits, horse or human). The polynucleotides and polypeptides are also useful as vaccines for treating or preventing these diseases. The polynucleotides are useful for gene therapy. The present sequence is that of one of a group of polynucleotides (ABQ88745-ABQ90015) disclosed electronically as sequences of the invention. However only 1271 polynucleotide sequences are given, whereas 1477 polynucleotides and 91 proteins are claimed.
CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequence.
XX
XX SQ Sequence 436 BP; 92 A; 89 C; 93 G; 129 T; 33 other;
Query Match 80.0%; Score 20; DB 24; Length 436;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTTCAGTGGCCGAGATA 20
|||||
Db 278 GGTTCAGTGGCCGAGATA 259
RESULT 41
AAK62995/c
ID AAK62995 standard; cDNA; 451 BP.
XX
XX AC AAK62995;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:8055.
DE

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ss.
XX Homo sapiens.
XX WO200157182-A2.
XX 09-AUG-2001.
XX 17-JAN-2001; 2001WO-US01354.
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 18-AUG-2000; 2000US-0225759.
PR 22-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0241826.
PR 08-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 05-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.

PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI: 2001-483426/52.
 DR P-PSDB; AAM90214.
 XX
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 XX Claim 1; SEQ ID NO 8055; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I). By inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 451 BP; 91 A; 93 C; 107 G; 154 T; 6 other;
 Query Match 80.0%; Score 20; DB 22; Length 451;
 Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTTCAGTGCAGCCGAGATA 20
 Db 284 GGTTCAGTGCAGCCGAGATA 265
 RESULT 42
 ABV45555/c
 ID ABV45555 standard; cDNA; 459 BP.
 XX
 AC ABV45555;
 XX
 DT 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 45546.
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US05171.
 XX
 PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 DR WPI: 2001-662795/76.
 XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX
 PS Claim 1; Page 9009; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV52213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX
 SQ Sequence 459 BP; 98 A; 113 C; 113 G; 135 T; 0 other;
 Query Match 80.0%; Score 20; DB 23; Length 459;
 Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGTTCAGTGCAGCCGAGATA 20
 Db 158 GGTTCAGTGCAGCCGAGATA 139
 RESULT 43
 ABV47512
 ID ABV47512 standard; cDNA; 472 BP.
 XX
 AC ABV47512;
 XX
 DT 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 47503.
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US05171.
 XX
 PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 DR WPI: 2001-662795/76.
 XX

PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX
 XX Claim 1; Page 9340-9341; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX
 XX Sequence 472 BP; 157 A; 116 C; 103 G; 96 T; 0 other;
 SQ
 Query Match 80.0%; Score 20; DB 23; Length 472;
 Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATA 20
 |||||
 Db 153 GGTTCAGTGAGCCGAGATA 172

RESULT 44
 ABV46458
 ID ABV46458 standard; cDNA; 486 BP.
 XX
 AC ABV46458;
 XX
 XX
 XX 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 46449.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US05171.
 XX
 XX 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX
 XX Schlegel R, Endege WO, Monahan JE;
 PI
 DR WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
 XX prostate cells and correlating with presence of prostate cancer, useful
 XX for detecting presence of prostate cancer, stage of prostate cancer -
 XX

PS Claim 1; Page 9166; 11750pp; English.

XX

CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX
 XX Sequence 486 BP; 167 A; 116 C; 108 G; 95 T; 0 other;
 SQ
 Query Match 80.0%; Score 20; DB 23; Length 486;
 Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATA 20
 |||||
 Db 257 GGTTCAGTGAGCCGAGATA 276

RESULT 45
 ABV42598
 ID ABV42598 standard; cDNA; 533 BP.
 XX
 AC ABV42598;
 XX
 XX 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 42589.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US05171.
 XX
 XX 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX
 XX Schlegel R, Endege WO, Monahan JE;
 PI
 DR WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
 XX prostate cells and correlating with presence of prostate cancer, useful
 XX for detecting presence of prostate cancer, stage of prostate cancer -
 XX
 XX Claim 1; Page 8524; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX

SQ Sequence 533 BP; 142 A; 127 C; 134 G; 129 T; 1 other;

Query Match 80.0%; Score 20; DB 23; Length 533;
 Best Local Similarity 100.0%; Pred. No. 0.02;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATA 20

|||||

Db 418 GGTTCAGTGAGCCGAGATA 437

Search completed: November 24, 2002, 13:16:32
 Job time : 254.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2002, 12:31:33 ; Search time 1611.5 Seconds
(without alignments)
451.486 Million cell updates/sec

Title: US-09-784-423-124
Perfect score: 25
Sequence: 1 GGTGTCAGTGGCCGAGATAAGAGT 25

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rnd.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	25	100.0	25	6	AR154074	AR154074 Sequence
2	25	100.0	1000	6	AR153982	AR153982 Sequence
3	25	100.0	69964	9	HS496C20	Z83847 Human DNA s
4	25	100.0	138459	2	AC090795	AC090795 Homo sapi
5	23	92.0	99035	9	HS90012	AL031506 Human DNA
6	22	88.0	65613	9	AL391644	AL391644 Human DNA
7	22	88.0	117824	9	AL389886	AL389886 Human DNA
8	22	88.0	126482	9	AL137220	AL137220 Human DNA
9	22	88.0	148396	9	AL157400	AL157400 Human DNA
10	22	88.0	148623	2	AC068582	AC068582 Homo sapi
11	22	88.0	164857	9	AL161448	AL161448 Human DNA
12	22	88.0	173645	2	AC010770	AC010770 Homo sapi
13	22	88.0	180315	2	AC068460	AC068460 Homo sapi
14	22	88.0	195068	9	AC012100	AC012100 Homo sapi
15	22	88.0	198141	9	AL355499	AL355499 Human DNA
16	22	88.0	209790	2	AL512373	AL512373 Homo sapi
17	21	84.0	355	6	AX070981	AX070981 Sequence
18	21	84.0	520	9	HSPA27G11	Z79279 H. sapiens f
19	21	84.0	608	6	AX390380	AX390380 Sequence
20	21	84.0	1242	6	AX195315	AX195315 Sequence
21	21	84.0	1838	9	HOMHER2B	M12036 Human tyros
22	21	84.0	2843	6	AX195330	AX195330 Sequence
23	21	84.0	3761	9	HSAC000981	AC000981 Homo sapi
24	21	84.0	16951	6	AX195368	AX195368 Sequence
25	21	84.0	16952	9	AY044229	AY044229 Homo sapi
26	21	84.0	24144	2	AC092149	AC092149 Homo sapi
27	21	84.0	25889	9	HSU112A12	Z68225 Human DNA s
28	21	84.0	26913	9	AC108170	AC108170 Homo sapi
29	21	84.0	31163	9	AL354801	AL354801 Human DNA
30	21	84.0	35581	9	AC005331	AC005331 Homo sapi
31	21	84.0	37448	9	AC006138	AC006138 Homo sapi
32	21	84.0	38429	9	AC000035	AC000035 Homo sapi
33	21	84.0	38540	9	HSU46H11	Z82254 Human DNA s
34	21	84.0	38634	9	AC093236	AC093236 Homo sapi
35	21	84.0	40014	9	AC011516	AC011516 Homo sapi
36	21	84.0	40943	9	HSB42B1	Z75890 Human DNA s
37	21	84.0	42304	9	AL355881	AL355881 Human DNA
38	21	84.0	42874	9	AL357952	AL357952 Human DNA
39	21	84.0	43456	9	AL357952	AL357952 Human DNA
40	21	84.0	44986	2	AC068187	AC068187 Homo sapi
41	21	84.0	46213	9	AC005387	AC005387 Homo sapi
42	21	84.0	49228	9	AB065918	AB065918 Homo sapi
43	21	84.0	49999	6	AX015915	AX015915 Sequence
44	21	84.0	52687	2	AC068390	AC068390 Homo sapi
45	21	84.0	59305	2	AC091585	AC091585 Homo sapi

ALIGNMENTS

RESULT 1
AR154074
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

AR154074
Sequence 124 from patent US 6238863.
AR154074
AR154074.1 GI:15122127
Unknown.
Unclassified.
1 (bases 1 to 25)
Schumm,J.W. and Bacher,J.W.
Materials and methods for indentifying and analyzing intermediate
tandem repeat DNA markers
Patent: US 6238863-A 124 29-MAY-2001;

25 bp DNA linear PAT 08-AUG-2001


```

/note="MER5A repeat: matches 57. .185 of consensus"
11303. .11539
/note="L2 repeat: matches 1865. .2127 of consensus"
13046. .13314
/note="AlusX repeat: matches 28. .305 of consensus"
13527. .13983
/note="L1MA8 repeat: matches 5393. .5854 of consensus"
13984. .14293
/note="AlusQ repeat: matches 1. .310 of consensus"
14294. .14347
/note="L1MA8 repeat: matches 5340. .5393 of consensus"
14348. .14649
/note="AlusQ repeat: matches 2. .304 of consensus"
14651. .14966
/note="AluY repeat: matches 1. .306 of consensus"
14967. .15007
/note="L1MA8 repeat: matches 5352. .5341 of consensus"
15024. .16371
/note="L1M4 repeat: matches 3784. .5097 of consensus"
16413. .16611
/note="L1 repeat: matches 3186. .3382 of consensus"
16662. .16875
/note="MIR repeat: matches 29. .256 of consensus"
16891. .17310
/note="MLT2FB repeat: matches 2. .414 of consensus"
17137. .17686
/note="match: GSS: Em:AQ551353"
17359. .17438
/note="MER91A repeat: matches 114. .193 of consensus"
17825. .18044
/note="LTR23 repeat: matches 210. .437 of consensus"
18046. .18357
/note="AlusQ repeat: matches 1. .309 of consensus"
18511. .18622
/note="AlusQ repeat: matches 1. .313 of consensus"
18830. .18897
/note="BURL repeat: matches 6300. .6367 of consensus"
19044. .19080
/note="MIR repeat: matches 207. .245 of consensus"
19081. .19391
/note="AlusQ repeat: matches 1. .307 of consensus"
19392. .19492
/note="MIR repeat: matches 117. .207 of consensus"
19581. .19801
/note="MIR repeat: matches 29. .261 of consensus"
20302. .20464
/note="AlusQ/x repeat: matches 134. .302 of consensus"
20465. .20747
/note="AlusQ repeat: matches 1. .296 of consensus"
22217. .22508
/note="AlusX repeat: matches 14. .305 of consensus"
22764. .22848
/note="MIR repeat: matches 164. .249 of consensus"
23209. .23725
/note="CpG island"
/evidence-not_experimental
23946. .24063
/note="MIR repeat: matches 21. .140 of consensus"
25053. .25142
/note="MIR repeat: matches 21. .119 of consensus"
25207. .25233
/note="9 copies 3 mer gct 96 conserved"
25401. .26038
/note="match: GSS: Em:B54652"
25877. .26155
/note="AluJb repeat: matches 1. .296 of consensus"
26997. .27304
/note="AluJo repeat: matches 1. .311 of consensus"
27326. .27595
/note="AlusX repeat: matches 37. .311 of consensus"
28098. .28406
/note="AlusQ repeat: matches 1. .311 of consensus"
28847. .29128

```

```

misc_feature
/note="L2 repeat: matches 2362. .2745 of consensus"
complement(29031. .29545)
/note="match: GSS: Em:AQ528941"
30940. .30989
/note="25 copies 2 mer gt 100 conserved"
30990. .31054
/note="MER41A repeat: matches 400. .464 of consensus"
31120. .31339
/note="match: STS: Em:G03731; match: STS: Em:G03731"
34416. .34721
/note="AluY repeat: matches 1. .310 of consensus"
34776. .34821
/note="L2 repeat: matches 2688. .2733 of consensus"
34944. .35243
/note="AlusX repeat: matches 1. .300 of consensus"
35351. .35551
/note="MER20 repeat: matches 1. .204 of consensus"
36460. .36544
/note="L2 repeat: matches 2665. .2749 of consensus"
37480. .37620
/note="MIR repeat: matches 4. .150 of consensus"
37861. .38188
/note="AluJo repeat: matches 1. .312 of consensus"
complement(38024. .38419)
/note="match: STS: Em:G28177; match: STS: Em:R54672"
complement(38086. .38419)
/note="match: STS: Em:G25989"
38802. .38948

```

```

Query Match 100.0%; Score 25; DB 9; Length 69964;
Best Local Similarity 100.0%; Pred. No. 9.6e-06;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGTTCAGTCAGCCGAGATAAGACT 25
|||||

```

```

Db 6244 GGTTCAGTCAGCCGAGATAAGACT 6220

```

RESULT 4

AC090795

LOCUS

DEFINITION

AC090795

AC090795

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

AC090795.2

Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A., and Zody, M.

TITLE JOURNAL

Direct Submission

Submitted (10-MAR-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 11, 2001 this sequence version replaced gi:13270664.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L12790

Center clone name: 254_K_5

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 133567 bases at least Q40

Consensus quality: 135811 bases at least Q30

Consensus quality: 136678 bases at least Q20

Insert size: 139000; agarose-fp

Quality coverage: 137359; sum-of-contigs

Quality coverage: 6.5 in Q20 bases; agarose-fp

Quality coverage: 6.6 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1411: contig of 1411 bp in length
1412 1511: gap of 100 bp
1512 2165: contig of 654 bp in length
2166 2265: gap of 100 bp
2266 3074: contig of 809 bp in length
3075 3174: gap of 100 bp
3175 4763: contig of 1589 bp in length
4764 4863: gap of 100 bp
4864 6161: contig of 1298 bp in length
6162 6261: gap of 100 bp
6262 8601: contig of 2340 bp in length
8602 8701: gap of 100 bp
8702 13139: contig of 4438 bp in length
13140 13239: gap of 100 bp
13240 17613: contig of 4374 bp in length
17614 17713: gap of 100 bp
17714 30894: contig of 13181 bp in length
30895 30994: gap of 100 bp
30995 40051: contig of 9057 bp in length
40052 40151: gap of 100 bp
40152 95859: contig of 55708 bp in length
95860 95959: gap of 100 bp
95960 138459: contig of 42500 bp in length.

FEATURES source

Location/Qualifiers
1..138459
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"

misc_feature
1..1411
/clone="RP11-254K5"
/clone_lib="RPC1-11 Human Male BAC"
/note="assembly_fragment
clone_end:SP6
vector_side:left"

misc_feature 1512..2165
/note="assembly_fragment"
misc_feature 2266..3074
/note="assembly_fragment"
misc_feature 3175..4763
/note="assembly_fragment"
misc_feature 4864..6161
/note="assembly_fragment"
misc_feature 6262..8601
/note="assembly_fragment"
misc_feature 8702..13139
/note="assembly_fragment"
misc_feature 13240..17613
/note="assembly_fragment"
misc_feature 17714..30894
/note="assembly_fragment"
misc_feature 30995..40051
/note="assembly_fragment"
misc_feature 40152..95859
/note="assembly_fragment"
misc_feature 95960..138459
/note="assembly_fragment
clone_end:T7
vector_side:right"
BASE COUNT 38140 a 31260 c 30041 g 37914 t 1104 others
ORIGIN

Query Match 100.0%; Score 25; DB 2; Length 138459;
Best Local Similarity 100.0%; Pred. No. 8.9e-06;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGATAAGAGT 25
|||||

Db 33960 GGTTCAGTGCAGCCGAGATAAGAGT 33984
|||||

RESULT 5

HS90012/c

LOCUS

DEFINITION

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Comment

HS90012 99035 bp DNA linear PRI 23-NOV-1999
Human DNA sequence from clone 90012 on chromosome 6p24.1-25.1.
Contains STSS and GSSs, complete sequence.

AL031906

HTG

Homo sapiens.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 99035)

Phillips, S

Direct Submission

Submitted (03-NOV-1998) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

IMPORTANT: This sequence is not the entire insert of clone 90012.

It may be shorter because we only sequence overlapping sections

once, or longer because we arrange for a small overlap between

neighbouring submissions.

The true left end of clone 103M22 (AL031904) is at 98936 in this

sequence.

This sequence has been finished according to sequence map criteria

as follows. An attempt is made to resolve all sequencing problems,

such as compressions and repeats, but not necessarily within known

annotated human repeat sequence elements (e.g. Alu). Where the

sequence is ambiguous, there is an annotation using the 'unsure'

feature key.

This sequence was generated from part of bacterial clone contigs of

human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6>
 90012 is from the library RPC11 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/VECTOR:PCYPAC2>.

FEATURES

Source

1..99035
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="6"
 /map="p24.1-25.1"
 /clone="RP1-90012"
 /clone_lib="RPC1-1"

repeat_region
 1..272
 /note="L1MA9 repeat: matches 5727..5999 of consensus"
 repeat_region
 641..888
 /note="MLT1C repeat: matches 219..466 of consensus"
 repeat_region
 893..1047
 /note="L2 repeat: matches 2213..2363 of consensus"
 repeat_region
 1048..1553
 /note="MER44B repeat: matches 40..533 of consensus"
 repeat_region
 1570..1918
 /note="L2 repeat: matches 2375..2750 of consensus"
 repeat_region
 1844..1990
 /note="MIR repeat: matches 100..262 of consensus"
 repeat_region
 2561..2586
 /note="L3 copies 2 mer aa 92% conserved"
 repeat_region
 2591..2707
 /note="FLAM_A repeat: matches 1..117 of consensus"
 repeat_region
 2817..3122
 /note="AluX repeat: matches 1..306 of consensus"
 repeat_region
 3870..4087
 /note="L2 repeat: matches 2410..2656 of consensus"
 repeat_region
 4731..5046
 /note="AluYb8 repeat: matches 1..310 of consensus"
 repeat_region
 5556..5865
 /note="AluX repeat: matches 1..310 of consensus"
 repeat_region
 5911..6209
 /note="AluY repeat: matches 1..299 of consensus"
 repeat_region
 6534..6834
 /note="AluSq repeat: matches 1..311 of consensus"
 repeat_region
 7502..7541
 /note="L20 copies 2 mer tt 80% conserved"
 repeat_region
 8939..8983
 /note="L15 copies 3 mer tot 91% conserved"
 repeat_region
 9212..9514
 /note="AluSq repeat: matches 1..311 of consensus"
 repeat_region
 10567..10654
 /note="MIR repeat: matches 78..170 of consensus"
 repeat_region
 10680..10719
 /note="L10 copies 4 mer tgg 88% conserved"
 repeat_region
 10680..10717
 /note="L9 copies 2 mer tg 90% conserved"
 repeat_region
 10900..11200
 /note="AluSq repeat: matches 3..303 of consensus"
 repeat_region
 11900..12156
 /note="AluYb repeat: matches 63..307 of consensus"
 repeat_region
 12197..12740
 /note="L2 repeat: matches 1811..2419 of consensus"
 repeat_region
 14423..14470
 /note="L24 copies 2 mer tg 88% conserved"
 repeat_region
 15868..16027
 /note="MIR repeat: matches 53..212 of consensus"
 misc_feature
 18445..18878
 /note="match: GSSs AQ124934 AQ064640"
 repeat_region
 20722..20783
 /note="L31 copies 2 mer ta 74% conserved"
 repeat_region
 20722..20781
 /note="L15 copies 4 mer tata 73% conserved"
 misc_feature
 complement(21988..22391)
 /note="match: GSS AQ030463"
 repeat_region
 22517..22625

/note="L1M4 repeat: matches -5..103 of consensus"
 23263..23549
 /note="AluYb repeat: matches 3..296 of consensus"
 24442..24732
 /note="AluYb repeat: matches 17..304 of consensus"
 25195..25233
 /note="L1PA7 repeat: matches 6107..6163 of consensus"
 27460..27736
 /note="AluSc repeat: matches 1..291 of consensus"
 28336..28518
 /note="L2 repeat: matches 2549..2750 of consensus"
 30201..30248
 /note="L24 copies 2 mer ta 81% conserved"
 30259..30310
 /note="L26 copies 2 mer ac 77% conserved"
 31062..31097
 /note="9 copies 4 mer ataa 83% conserved"
 31099..31208
 /note="MIR repeat: matches 1..115 of consensus"
 31752..31855
 /note="L52 copies 2 mer aa 63% conserved"
 31887..32169
 /note="AluY repeat: matches 1..283 of consensus"
 32431..32603
 /note="L1M4 repeat: matches 5114..5298 of consensus"
 33650..34808
 /note="L2 repeat: matches 408..1636 of consensus"
 34822..35129
 /note="MLT1A1 repeat: matches 1..307 of consensus"
 35138..36578
 /note="L1PA2 repeat: matches 4712..6146 of consensus"
 37487..37671
 /note="MLT11 repeat: matches 167..352 of consensus"
 37793..38386
 /note="L2 repeat: matches 1911..2527 of consensus"
 39061..39184
 /note="L1M4 repeat: matches 2876..2991 of consensus"
 39577..39762
 /note="MER58A repeat: matches 7..207 of consensus"
 42238..43218
 /note="L2 repeat: matches 1326..2404 of consensus"
 43305..43446
 /note="L2 repeat: matches 2550..2735 of consensus"
 complement(44389..44852)
 /note="match: GSS AQ170695"
 complement(45574..46231)
 /note="match: GSS B49432"
 46399..46690
 /note="AluYb repeat: matches 3..308 of consensus"
 47337..47612
 /note="AluSc repeat: matches 1..277 of consensus"
 47427..47862
 /note="match: SRS AL031017"
 47863..47980
 /note="MIR repeat: matches 74..185 of consensus"
 47981..48423
 /note="MLT1C repeat: matches 1..464 of consensus"
 48424..48464
 /note="MIR repeat: matches 33..74 of consensus"
 49974..50286
 /note="AluX repeat: matches 1..311 of consensus"
 complement(50053..50357)
 /note="match: GSS AQ082392"
 51121..51355
 /note="AluSq repeat: matches 1..236 of consensus"
 52036..52753
 /note="L1PA2 repeat: matches 5430..6146 of consensus"
 52795..52854
 /note="HAL1 repeat: matches 1552..1605 of consensus"
 54067..54128
 /note="MER94 repeat: matches 66..131 of consensus"
 54129..54420
 /note="AluYb repeat: matches 6..298 of consensus"

```

repeat_region 54421..54498
/note="MER94 repeat: matches 1..66 of consensus"
55613..55675
/note="9 copies 7 mer aatatat 76% conserved"
55809..55990
/note="26 copies 7 mer attatat 64% conserved"
56456..56626
/note="FRAM repeat: matches -6..163 of consensus"
57248..57293
/note="MER70A repeat: matches 544..585 of consensus"
57294..57570
/note="AluX repeat: matches 17..299 of consensus"
57571..57945
/note="MER70A repeat: matches 134..544 of consensus"
59109..59411
/note="AluX repeat: matches 1..304 of consensus"
60599..60628
/note="15 copies 2 mer tt 100% conserved"
62412..62449

Query Match 92.0%; Score 23; DB 9; Length 99035;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCGAGATAGA 23
|||||
DB 51139 GGTTCAGTGCAGCGAGATAGA 51117

RESULT 6
AL391644
LOCUS
DEFINITION Human DNA sequence from clone RP11-535F17 on chromosome 6 Contains
part of a KH domain protein; similar to mammalian Sam68, ESTs and
GSSs, complete sequence.
ACCESSION AL391644
VERSION AL391644.12 GI:11229206
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Bates, K.
Direct Submission
Submitted (09-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Nov 20, 2000 this sequence version replaced gi:11190597.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WormPEP; Information
on the WormPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
IMPORTANT: This sequence is not the entire insert of clone
RP11-535F17. It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true left end of clone RP3-522D12 is at 65514 in this sequence.
The true right end of clone RP1-240B8 is at 100 in this sequence.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such

```

as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP11-535F17 is from the library RPI-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6.

FEATURES	Location/Qualifiers
source	1..65613
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="6"
	/clone="RP11-535F17"
	/clone_lib="RPI-11.2"
	739..1206
repeat_region	/note="L2 repeat: matches 2266..2745 of consensus"
repeat_region	2904..3068
misc_feature	/note="MIR repeat: matches 80..259 of consensus"
	3900..4345
misc_feature	/note="match: GSS: Em:AQ731895"
	3904..4430
misc_feature	/note="match: GSS: Em:AQ489033"
	3908..4305
repeat_region	/note="match: GSS: Em:AQ566108"
	4438..4626
repeat_region	/note="L2 repeat: matches 2273..2470 of consensus"
	4794..4938
misc_feature	/note="MER5A repeat: matches 29..182 of consensus"
	5354..5758
repeat_region	/note="match: GSS: Em:B37233"
	6506..6709
repeat_region	/note="MIR repeat: matches 31..262 of consensus"
	6831..7153
repeat_region	/note="MER46C repeat: matches 1..337 of consensus"
	7475..7602
repeat_region	/note="L1PB3 repeat: matches 4636..4756 of consensus"
	7592..8952
repeat_region	/note="L1PB3 repeat: matches 4748..6244 of consensus"
	9957..10037
repeat_region	/note="MLT1F repeat: matches 83..163 of consensus"
	10008..10435
misc_feature	/note="MLT1F repeat: matches 122..535 of consensus"
	complement(10824..11284)
repeat_region	/note="match: GSS: Em:AQ784097"
	11909..12105
repeat_region	/note="MIR repeat: matches 49..261 of consensus"
	12406..12626
repeat_region	/note="LTR33 repeat: matches 1..228 of consensus"
	13729..16230
misc_feature	/note="L1PA3 repeat: matches 3643..6146 of consensus"
	16152..16296
repeat_region	/note="match: GSS: Em:AQ078529"
	16652..16943
repeat_region	/note="AluX repeat: matches 8..299 of consensus"
	16945..17436
repeat_region	/note="L1PA11 repeat: matches 773..1252 of consensus"
	17451..17510
repeat_region	/note="AluSc repeat: matches 240..299 of consensus"
	17525..22399
repeat_region	/note="L1PA11 repeat: matches 1246..6164 of consensus"
	22792..22825
repeat_region	/note="L1PA13 repeat: matches 6111..6144 of consensus"
	22827..23369
repeat_region	/note="L1PA12 repeat: matches 5550..6100 of consensus"
	23371..23909
repeat_region	/note="L1ME repeat: matches 5132..5668 of consensus"
	24061..24150
repeat_region	/note="L2 repeat: matches 2620..2710 of consensus"
	24570..24687
repeat_region	/note="L2 repeat: matches 2373..2489 of consensus"
	25807..26027
gene	/note="L1MC4 repeat: matches 5861..6116 of consensus"
	complement(26624..26706)


```

source
1..117824
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/clone="RPI-237C24"
/clone_lib="RPCI-1"
1..65
/note="AluSc repeat: matches 232..296 of consensus"
/note="L1 repeat: matches 4731..4854 of consensus"
346..641
/note="AluSx repeat: matches 1..308 of consensus"
978..1271
/note="AluJb repeat: matches 1..289 of consensus"
1273..1382
/note="Tandem repeat. Tandem repeat shows differences
between individual subclones. Overall assembly consistent
with restriction digest."
1275..1379
/note="3 copies 35 mer 78% conserved"
1276..1381
/note="53 copies 2 mer 99 67% conserved"
1288..1386
/note="3 copies 33 mer 78% conserved"
1629..1845
/note="AluJb repeat: matches 2..219 of consensus"
2306..2407
/note="L1Pall1 repeat: matches 6064..6164 of consensus"
2654..2727
/note="L1M2 repeat: matches 5539..5594 of consensus"
2728..3037
/note="AluSx repeat: matches 1..310 of consensus"
3038..3088
/note="L1M2 repeat: matches 5594..5649 of consensus"
4213..4471
/note="MER46A repeat: matches 1..236 of consensus"
4484..4649
/note="Tigger3(Golem) repeat: matches 2833..3028 of
consensus"
4650..5005
/note="MER1B repeat: matches 1..337 of consensus"
5006..5039
/note="Tigger3(Golem) repeat: matches 2804..2833 of
consensus"
5040..5228
/note="AluJb repeat: matches 113..301 of consensus"
5313..5633
/note="AluJb repeat: matches 5..308 of consensus"
6563..6864
/note="AluI repeat: matches 1..301 of consensus"
7151..7256
/note="MER30 repeat: matches 3..96 of consensus"
7257..7563
/note="AluSx repeat: matches 1..308 of consensus"
7564..7661
/note="MER30 repeat: matches 96..230 of consensus"
8150..8215
/note="MER2 repeat: matches 268..342 of consensus"
8246..8373
/note="MER2 repeat: matches 88..210 of consensus"
8414..8715
/note="AluJc repeat: matches 1..294 of consensus"
8717..8811
/note="MER2 repeat: matches 2..97 of consensus"
complement(8858..89217)
/gene="SLC23A1"
complement(join(8858..13705,15790..15885,18449..18586,
19279..19408,24271..24376,26407..26554,30437..30593,
31077..31197,40143..40324,41254..41324,42322..42410,
56056..56213,58943..59059,69381..69479,88956..89217))
/gene="SLC23A1"
/product="dJ237C24.1 (Solute carrier family 23 (nucleobase
transporters), member 1 (KIAA0238))"

polyA_site
polyA_signal
repeat_region
CDS
/note="13 copies 2 mer 99 100% conserved"
complement(join(13473..13705,15790..15885,18449..18586,
19279..19408,24271..24376,26407..26554,30437..30593,
31077..31197,40143..40324,41254..41324,42322..42410,
56056..56213,58943..59059,69381..69479,88956..89063))
/gene="SLC23A1"
/note="match; proteins: Tr:Q9Z2J0 Tr:Q9UI39 Tr:Q9UH7
Tr:Q95191 Tr:Q9WTW8 Tr:Q9WTW7"
/codon_start=1
/evidence-not_experimental
/product="dJ237C24.1 (Solute carrier family 23 (nucleobase
transporters), member 1 (KIAA0238))"
/protein_id="CAC16126.1"
/db_xref="GI:12314277"
/translation="MMGIGKNTTSKMEAGSTGEYDEAKHPAFFTPVWINGGAT
SSGQDNDETELMATYTGIEAGSLAETLDSGLDQPSDMITTYEDVPPWYLC
IFLQGHVYFCFSGTAYVPELLADAMCVGDWATSLIGTIFFCVGLITLQTFGC
RLPQAFAPFATAPARAILSLDKWCTNTDYSVANGTAELTHEWPRIRIQGA
IIMSLLEVIGLIGPALLKYGITTPYVALLGSLFGQAGERAGKHWGJAMUT
IFVLFSOYARNVKPEPLPYTKSKGWTAKQLFKMFPIALILVSWLLCFITVD
VFPDSTKGYFARTDAQOGLVAVFWKFPYPPFOWGLPTVSAAGVIGMLSAVASTI
ESIDYFACARLSCAPPPHAINRGIFVEGLSCVLDIGFTGNGSTSSNIGVLGI
TKGSRVTCQCALMALMGKIFGALFASLPDPLVCALECFGLMTAVLGNLOF
IDLNSRNLFVLGFIFFGLPSYLRONPLVTGIGTDQVLNLLTAMFVGGCVAF
ILDNTIPGTPEERGIKWKGVKGNKSLDMESYNLPFGWNIITIKKRCFLYLPISPT
FVGTWKGLRKSDNSRSDSDSQATG"
13776..13837
/note="31 copies 2 mer ca 96% conserved"
13779..13846
/note="2 copies 34 mer 91% conserved"
14150..14460
/note="AluSx repeat: matches 1..311 of consensus"
14471..14538
/note="2 copies 34 mer 92% conserved"
16112..16830
/note="match; GSS: Em:AQ319998"
16466..16649
/note="L1MB5 repeat: matches 5482..5667 of consensus"
16674..17043
/note="MER82 repeat: matches 228..641 of consensus"
17040..17129
/note="MER82 repeat: matches 1..97 of consensus"
17138..17259
/note="L1MB5 repeat: matches 5705..5822 of consensus"
17260..17552
/note="AluSx repeat: matches 1..289 of consensus"
17553..17873
/note="L1MB5 repeat: matches 5822..6161 of consensus"
19877..20061
/note="MIR repeat: matches 28..238 of consensus"
21206..21582
/note="match; GSS: Em:AQ603395"
21270..21381
/note="L1 repeat: matches 2337..2459 of consensus"
21901..21977
/note="L1 repeat: matches 2675..2750 of consensus"
21920..21986
/note="MIR repeat: matches 184..251 of consensus"

```

```

repeat_region 22173..22329
/note="MER2 repeat: matches 1..171 of consensus"
repeat_region 22330..22633
/note="AluX repeat: matches 1..304 of consensus"
repeat_region 22634..22854

```

```

Query Match 88.0%; Score 22; DB 9; Length 117824;
Best Local Similarity 100.0%; Pred.No.0.00079;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGTGTCAGTGACCGAGATAAG 22
      |||
Db 48770 GGTGTCAGTGACCGAGATAAG 48749

```

```

RESULT 8
AL137220/c
LOCUS
DEFINITION
Human DNA sequence from clone RP11-354I10 on chromosome 6 Contains
GSSs and STSs, complete sequence.
ACCESSION
AL137220
VERSION
AL137220.8 GI:9367921
KEYWORDS
HTG.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 126482)
Tromans,A.
Direct Submission
Submitted (07-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
Clone
Requests: clonerequests@sanger.ac.uk
On Jul 22, 2000 this sequence version replaced gi:9366903.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; WP., WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
RP11-354I10 is from the library RPCI-11.2 constructed at the
Roswell Park Cancer Institute by the group of Pieter de Jong. For
further details see http://bacpac.med.buffalo.edu/
VECTOR: pBACE3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-354I10 It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true right end of clone RP11-354I10 is at 126482 in this
sequence. The true left end of clone RP1-246P14 is at 103489 in
this sequence. The true right end of clone RP1-133H11 is at 100 in
this sequence.

```

```

FEATURES
            Location/Qualifiers
             i. 126482
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /chromosome="6"
              /clone="RP11-354I10"
              /clone_lib="RPCI-11.2"

```

```

repeat_region 10..103
/note="MSTD repeat: matches 76..194 of consensus"
153..280
/note="MSTD repeat: matches 283..393 of consensus"
767..1065
/note="AluJo repeat: matches 1..310 of consensus"
1123..1225
/note="L2 repeat: matches 2636..2750 of consensus"
1981..2306
/note="AluSp repeat: matches 1..313 of consensus"
2483..2843
/note="L2 repeat: matches 10..410 of consensus"
2871..3017
/note="AluSg/x repeat: matches 166..312 of consensus"
3018..4467
/note="PTRS repeat: matches 743..2438 of consensus"
3046..3514
/note="CpG island"
/evidence=not_experimental
4800..5290
/note="MLTIF repeat: matches 3..534 of consensus"
5426..5597
/note="L2 repeat: matches 1282..1453 of consensus"
5624..5834
/note="L1PB2 repeat: matches 5938..6155 of consensus"
5835..6125
/note="AluSp repeat: matches 1..294 of consensus"
complement(6533..6889)
/note="match: GSS: Em:AQ726655"
6572..6650
/note="AluSg/x repeat: matches 210..288 of consensus"
8860..9098
/note="L2 repeat: matches 2455..2697 of consensus"
9501..9549
/note="MIR repeat: matches 91..139 of consensus"
9570..9875
/note="AluJb repeat: matches 1..308 of consensus"
complement(9694..10158)
/note="match: GSS: Em:B92811"
10254..10534
/note="MLT1B repeat: matches 152..390 of consensus"
10535..10844
/note="AluSg repeat: matches 1..309 of consensus"
10845..10974
/note="MLT1B repeat: matches 24..153 of consensus"
11330..11420
/note="MIR repeat: matches 72..176 of consensus"
11483..11767
/note="AluSc repeat: matches 1..285 of consensus"
14000..14186
/note="L1PAL6 repeat: matches 5967..6157 of consensus"
15407..15460
/note="MIR repeat: matches 77..139 of consensus"
16339..17122
/note="L1PAL3 repeat: matches 5372..6155 of consensus"
18620..18716
/note="L1 repeat: matches 3632..3722 of consensus"
18794..19034
/note="MIR repeat: matches 13..251 of consensus"
19102..19439
/note="LTR16A repeat: matches 92..445 of consensus"
20045..20483
/note="match: GSS: Em:AQ735674"
21172..21325
/note="L2 repeat: matches 2596..2750 of consensus"
22699..23307
/note="L1MEC repeat: matches 2063..2354 of consensus"
23381..23596
/note="L1M1 repeat: matches 1411..1613 of consensus"
23912..25032
/note="L1M1 repeat: matches 272..1418 of consensus"
24730..25141
/note="match: GSS: Em:AQ115747"

```

```

repeat_region 25372..25631
/note="AluI repeat: matches 38..299 of consensus"
27712..27759
/note="12 copies 4 mer taca 95% conserved"
27803..28296
/note="L2 repeat: matches 2059..2582 of consensus"
28379..28579
/note="MER46C repeat: matches 3..220 of consensus"
28863..29163
/note="AluSg repeat: matches 1..311 of consensus"
30914..31284
/note="MLTIB repeat: matches 1..387 of consensus"
31397..31464
/note="34 copies 2 mer tg 91% conserved"
complement(31554..32145)
/note="match: GSS: Em:AQ470977"
33808..34126
/note="MER7A repeat: matches 1..334 of consensus"
36084..36393
/note="AluSg repeat: matches 2..307 of consensus"
37452..37491
/note="L1M4c repeat: matches 1877..1918 of consensus"
37492..37801
/note="AluY repeat: matches 1..311 of consensus"
37802..37914
/note="L1M4c repeat: matches 1776..1877 of consensus"
38063..38724
/note="L1M4c repeat: matches 1037..1697 of consensus"
38722..38899
/note="L1MD repeat: matches 8..171 of consensus"
38996..39201
/note="L1M4 repeat: matches 2148..2362 of consensus"
39342..39525
/note="92 copies 2 mer ta 67% conserved"
39346..39525
/note="30 copies 6 mer tatata 67% conserved"
39721..39812
/note="46 copies 2 mer tt 62% conserved"
41159..41390
/note="Alu8g repeat: matches 1..234 of consensus"
41986..42075
/note="45 copies 2 mer tt 62% conserved"
complement(42516..42963)
/note="match: GSS: Em:AQ813487"
42837..42972
/note="match: SRS: Em:HSGBDSTS9"
42972..43671
/note="match: GSS: Em:AQ626117"
44171..44206
/note="9 copies 4 mer tgtg 83% conserved"
45154..45250
/note="LTR29 repeat: matches 4..100 of consensus"
45346..45908
/note="Tigger4(Zombi) repeat: matches 2167..2731 of consensus"
45909..46008
/note="50 copies 2 mer ac 78% conserved"
46009..46298
/note="AluJb repeat: matches 5..302 of consensus"
46740..46828
/note="MER39b repeat: matches 483..568 of consensus"
46810..46882
/note="LTR29 repeat: matches 555..617 of consensus"
48034..48168
/note="MIR repeat: matches 134..262 of consensus"
48891..49394
/note="match: GSS: Em:AQ720319"
50724..50791
/note="17 copies 4 mer tctc 73% conserved"
50893..50958
/note="MIR repeat: matches 62..145 of consensus"

Query Match 88.0%; Score 22; DB 9; Length 126482;

```

```

Best Local Similarity 100.0%; Pred. No. 0.00078;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTTCAGTGGCCGAGATAAG 22
|||||
Db 41177 GGTTCAGTGGCCGAGATAAG 41156
|||||

```

```

RESULT 9
AL157400/c
LOCUS Human DNA sequence from clone RP11-80H5 on chromosome 10, complete
sequence.
DEFINITION AL157400
ACCSSION AL157400
VERSION AL157400.8 GI:14596274
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens

```

```

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 148396)
JOURNAL Direct Submission

```

```

COMMENT Submitted (03-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jul 4, 2001 this sequence version replaced gi:14529754.
During difference assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em., EMBL; Sw.,
SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep

```

```

This sequence is the entire insert of clone RP11-80H5. The true left
end of clone RP11-248C1 is at 129612 in this sequence. The true
right end of clone RP11-9C23 is at 17504 in this sequence.

```

```

This sequence is the entire insert of clone RP11-80H5. The true left
end of clone RP11-248C1 is at 129612 in this sequence. The true
right end of clone RP11-9C23 is at 17504 in this sequence.

```

```

This sequence is the entire insert of clone RP11-80H5. The true left
end of clone RP11-248C1 is at 129612 in this sequence. The true
right end of clone RP11-9C23 is at 17504 in this sequence.

```

```

This sequence is the entire insert of clone RP11-80H5. The true left
end of clone RP11-248C1 is at 129612 in this sequence. The true
right end of clone RP11-9C23 is at 17504 in this sequence.

```

```

This sequence is the entire insert of clone RP11-80H5. The true left
end of clone RP11-248C1 is at 129612 in this sequence. The true
right end of clone RP11-9C23 is at 17504 in this sequence.

```

```

This sequence is the entire insert of clone RP11-80H5. The true left
end of clone RP11-248C1 is at 129612 in this sequence. The true
right end of clone RP11-9C23 is at 17504 in this sequence.

```

```

This sequence is the entire insert of clone RP11-80H5. The true left
end of clone RP11-248C1 is at 129612 in this sequence. The true
right end of clone RP11-9C23 is at 17504 in this sequence.

```

```

This sequence is the entire insert of clone RP11-80H5. The true left
end of clone RP11-248C1 is at 129612 in this sequence. The true
right end of clone RP11-9C23 is at 17504 in this sequence.

```

```

This sequence is the entire insert of clone RP11-80H5. The true left
end of clone RP11-248C1 is at 129612 in this sequence. The true
right end of clone RP11-9C23 is at 17504 in this sequence.

```

repeat_region 4493. 4609 /note="FLAM_A repeat: matches 11. .126 of consensus"
repeat_region 5205. 5373 /note="MER5A repeat: matches 8. .179 of consensus"
repeat_region 5712. 6009 /note="AluX repeat: matches 1. .296 of consensus"
repeat_region 6279. 6483 /note="AluJ repeat: matches 38. .256 of consensus"
repeat_region 6722. 7032 /note="AluX repeat: matches 1. .309 of consensus"
repeat_region 7526. 7837 /note="AluY repeat: matches 3. .311 of consensus"
repeat_region 8440. 8740 /note="AluY repeat: matches 2. .308 of consensus"
repeat_region 10302. 12532 /note="TIGGR1 repeat: matches 26. .2352 of consensus"
repeat_region 12541. 13173 /note="L1MB8 repeat: matches 5499. .6158 of consensus"
repeat_region 13172. 13391 /note="HALL repeat: matches 704. .924 of consensus"
repeat_region 14180. 14401 /note="L2 repeat: matches 2514. .2734 of consensus"
repeat_region 14825. 15123 /note="L2 repeat: matches 2170. .2486 of consensus"
repeat_region 15180. 15572 /note="L2 repeat: matches 2246. .2681 of consensus"
repeat_region 15829. 15882 /note="MER46C repeat: matches 280. .334 of consensus"
repeat_region 15885. 16187 /note="AluY repeat: matches 1. .309 of consensus"
repeat_region 16578. 16707 /note="MIR repeat: matches 19. .156 of consensus"
repeat_region 19115. 19162 /note="MIR repeat: matches 183. .231 of consensus"
repeat_region 19495. 19861 /note="MLT1A repeat: matches 1. .365 of consensus"
repeat_region 19994. 20170 /note="MIR repeat: matches 67. .251 of consensus"
repeat_region 20703. 20745 /note="MER53 repeat: matches 1. .43 of consensus"
repeat_region 21233. 21612 /note="AluX repeat: matches 5. .298 of consensus"
repeat_region 22283. 22903 /note="L1MA7 repeat: matches 5542. .6279 of consensus"
repeat_region 23700. 23756 /note="L2 repeat: matches 2641. .2701 of consensus"
repeat_region 24503. 24602 /note="50 copies 2 mer aa 68% conserved"
repeat_region 24889. 25028 /note="L1ME repeat: matches 5509. .5651 of consensus"
repeat_region 25288. 25392 /note="L1MC5 repeat: matches 7711. .7809 of consensus"
repeat_region 26800. 26938 /note="L1ME3A repeat: matches 6000. .6145 of consensus"
repeat_region 29645. 29947 /note="AluX repeat: matches 3. .301 of consensus"
repeat_region 30012. 30308 /note="AluX repeat: matches 1. .300 of consensus"
repeat_region 30309. 30412 /note="4 copies 26 mer 76% conserved"
repeat_region 32503. 32614 /note="MER20 repeat: matches 97. .218 of consensus"
repeat_region 34233. 34676 /note="MLT1C repeat: matches 1. .464 of consensus"
repeat_region 34886. 35055 /note="MIR repeat: matches 97. .259 of consensus"
repeat_region 35362. 35517 /note="MIR repeat: matches 72. .261 of consensus"
repeat_region 35671. 35972 /note="AluJ repeat: matches 5. .300 of consensus"
repeat_region 37601. 37801 /note="MER63A repeat: matches 5. .210 of consensus"
repeat_region 37986. 38081

repeat_region 38177. 38258 /note="MIR repeat: matches 70. .164 of consensus"
repeat_region 38765. 38887 /note="MER5A repeat: matches 109. .189 of consensus"
repeat_region 38888. 39187 /note="TIGGR1 repeat: matches 2295. .2418 of consensus"
repeat_region 39188. 39236 /note="AluX repeat: matches 4. .303 of consensus"
repeat_region 39237. 39524 /note="TIGGR1 repeat: matches 2247. .2295 of consensus"
repeat_region 39525. 40995 /note="AluY repeat: matches 1. .287 of consensus"
repeat_region 40996. 41297 /note="TIGGR1 repeat: matches 823. .2247 of consensus"
repeat_region 41298. 42061 /note="AluJ repeat: matches 1. .302 of consensus"
repeat_region 42505. 42804 /note="TIGGR1 repeat: matches 48. .823 of consensus"
repeat_region 43354. 43546 /note="AluY repeat: matches 1. .300 of consensus"
repeat_region 43547. 43838 /note="L2 repeat: matches 2364. .2534 of consensus"
repeat_region 43839. 43953 /note="AluX repeat: matches 1. .292 of consensus"
repeat_region 44660. 44762 /note="L2 repeat: matches 2534. .2707 of consensus"
repeat_region 45925. 46088 /note="MIR repeat: matches 47. .149 of consensus"
repeat_region 47234. 47399 /note="MIR repeat: matches 75. .261 of consensus"
repeat_region 48990. 49282 /note="MER5B repeat: matches 1. .177 of consensus"
repeat_region 50066. 50149 /note="AluJ repeat: matches 1. .292 of consensus"
repeat_region 50966. 51115 /note="AluJ repeat: matches 23. .111 of consensus"
repeat_region 51118. 51401 /note="FAM repeat: matches 20. .169 of consensus"
repeat_region 51558. 51697 /note="AluX repeat: matches 4. .287 of consensus"
repeat_region 53048. 53454 /note="MLT1 repeat: matches 291. .408 of consensus"
repeat_region 54338. 55405 /note="MIR repeat: matches 8. .410 of consensus"
repeat_region 55494. 55792 /note="MER1C repeat: matches 1. .1071 of consensus"
repeat_region 55976. 56049 /note="AluSP repeat: matches 1. .299 of consensus"
repeat_region 56206. 56487 /note="MIR repeat: matches 63. .139 of consensus"
repeat_region 57028. 57324 /note="AluX repeat: matches 1. .282 of consensus"
repeat_region 57685. 57996 /note="AluY repeat: matches 1. .295 of consensus"
repeat_region 58688. 58938 /note="AluY repeat: matches 1. .310 of consensus"
repeat_region 59164. 59333 /note="ORSL repeat: matches 222. .472 of consensus"
repeat_region 59378. 59674 /note="MER5A repeat: matches 2. .186 of consensus"
repeat_region 60778. 61062 /note="AluJ repeat: matches 1. .297 of consensus"
repeat_region 61641. 61965 /note="AluY repeat: matches 16. .300 of consensus"

Query Match 88.0%; Score 22; DB 9; Length 148396;
Best Local Similarity 100.0%; Pred. No. 0.00077;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCGGAGATAAG 22
|||||

Db 96484 GGTTCAGTGAGCGGAGATAAG 96463

RESULT 10

AC068582/c
 LOCUS
 DEFINITION Homo sapiens clone RP11-27C24, WORKING DRAFT SEQUENCE, 22 unordered
 pieces.
 AC068582
 VERSION AC068582.2 GI:8705085
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens

REFERENCE

AUTHORS
 TITLE 1 (bases 1 to 148623)
 JOURNAL
 REFERENCE
 AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
 Boguslavskiy,L., Bouckgalter,B., Brown,A., Burkett,G.,
 Campopiano,A., Castle,A., Choquel,Y., Colangelo,M., Collins,S.,
 Dodde,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
 Galagan,J., Gardyna,S., Ginde,S., Govette,M., Graham,L.,
 Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
 Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczyk,J.,
 Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
 McCarthy,M., McEwan,P., McKurk,A., McKernan,K., McPheeters,R.,
 Melidini,J., Meneus,L., Mihova,T., Miranda,C., Miengra,V., Morrow,J.,
 Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
 Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.
 Direct Submision
 Submitted (04-MAY-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jun 25, 2000 this sequence version replaced gi:7705139.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: L4683

Center clone name: 27_C_24

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 136479 bases at least Q40
 Consensus quality: 142634 bases at least Q30
 Consensus quality: 144995 bases at least Q20
 Insert size: 167000; agarose-ff
 Quality coverage: 4.2 in Q20 bases; agarose-ff
 Quality coverage: 4.8 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 22 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will

* be preserved.
 * 1 1594: contig of 1594 bp in length
 * 1595 1694: gap of 100 bp
 * 1695 2745: contig of 1051 bp in length
 * 2746 2845: gap of 100 bp
 * 2846 5356: contig of 2511 bp in length
 * 5357 5456: gap of 100 bp
 * 5457 8179: contig of 2723 bp in length
 * 8180 8279: gap of 100 bp
 * 8280 11291: contig of 3012 bp in length
 * 11292 11391: gap of 100 bp
 * 11392 13945: contig of 2554 bp in length
 * 13946 14045: gap of 100 bp
 * 14046 17830: contig of 3785 bp in length
 * 17831 17930: gap of 100 bp
 * 17931 21469: contig of 3539 bp in length
 * 21470 21569: gap of 100 bp
 * 21570 24779: contig of 3210 bp in length
 * 24780 24879: gap of 100 bp
 * 24880 27330: contig of 2451 bp in length
 * 27331 27430: gap of 100 bp
 * 27431 30523: contig of 3093 bp in length
 * 30524 30623: gap of 100 bp
 * 30624 34816: contig of 4193 bp in length
 * 34817 34916: gap of 100 bp
 * 34917 39866: contig of 4950 bp in length
 * 39867 39966: gap of 100 bp
 * 39967 45579: contig of 6613 bp in length
 * 45580 46679: gap of 100 bp
 * 46680 52861: contig of 6182 bp in length
 * 52862 52961: gap of 100 bp
 * 52962 61332: contig of 8371 bp in length
 * 61333 61432: gap of 100 bp
 * 61433 69233: contig of 7807 bp in length
 * 69240 69339: gap of 100 bp
 * 69340 82363: contig of 13024 bp in length
 * 82364 82463: gap of 100 bp
 * 82464 94321: contig of 11858 bp in length
 * 94322 94421: gap of 100 bp
 * 94422 108651: contig of 14230 bp in length
 * 108652 108751: gap of 100 bp
 * 108752 128711: contig of 19960 bp in length
 * 128712 128811: gap of 100 bp
 * 128812 148623: contig of 19812 bp in length.

FEATURES

source

1..148623
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="RP11-27C24"
 /clone_lib="RPC1-11 Human Male BAC"
 1..1594
 /note="assembly_fragment"
 1695..2745
 /note="assembly_fragment"
 clone_end:77
 vector_side:right
 2846..5356
 /note="assembly_fragment"
 5457..8179
 /note="assembly_fragment"
 8280..11291
 /note="assembly_fragment"
 11392..13945
 /note="assembly_fragment"
 14046..17830
 /note="assembly_fragment"
 17931..21469
 /note="assembly_fragment"
 21570..24779
 /note="assembly_fragment"
 24880..27330
 /note="assembly_fragment"
 27431..30523

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature 30624..34816 /note="assembly_fragment"
 misc_feature 34917..39866 /note="assembly_fragment"
 misc_feature 39967..46579 /note="assembly_fragment"
 misc_feature 46680..52861 /note="assembly_fragment"
 clone_end:SP6
 vector_side:right
 misc_feature 52962..61332 /note="assembly_fragment"
 misc_feature 61433..69239 /note="assembly_fragment"
 misc_feature 69340..82363 /note="assembly_fragment"
 misc_feature 82464..94321 /note="assembly_fragment"
 misc_feature 94422..108651 /note="assembly_fragment"
 misc_feature 108752..128711 /note="assembly_fragment"
 misc_feature 128812..148623 /note="assembly_fragment"
 BASE COUNT 42327 a 33848 c 32549 g 37797 t 2102 others
 ORIGIN

Query Match 88.0%; Score 22; DB 2; Length 148623;
 Best Local Similarity 100.0%; Pred. No. 0.00077;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTTGCAGTGGCGGAGATAAG 22
 |||||
 Db 134898 GTTGCAGTGGCGGAGATAAG 134877

RESULT 11
 AC010770
 LOCUS
 DEFINITION Human DNA sequence from clone RP11-381H12 on chromosome 9, complete sequence.
 ACCESSION AL161448
 VERSION AL161448.16 GI:18491317
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Tronans,A.
 Direct Submission
 Submitted (01-FEB-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 On Feb 4, 2002 this sequence version replaced gi:119339694.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr9
 RP11-381H12 is from the library RPI1-11.2 constructed by the group of Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pBacE3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-381H12. It may be shorter because we sequence overlapping sections only once, except for a short overlap.
 The true right end of clone RP11-381H12 is at 164857 in this sequence. The true left end of clone RP11-310F24 is at 148729 in this sequence. The true right end of clone RP11-376F21 is at 2000 in this sequence.

FEATURES
 Source
 1..164857
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="9"
 /clone="RP11-381H12"
 /clone_lib="RPI1-11.2"
 BASE COUNT 42782 a 40111 c 39202 g 42762 t
 ORIGIN

Query Match 88.0%; Score 22; DB 9; Length 164857;
 Best Local Similarity 100.0%; Pred. No. 0.00076;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GTTGCAGTGGCGGAGATAAGA 23
 |||||
 Db 56802 GTTGCAGTGGCGGAGATAAGA 56823

RESULT 12
 AC010770
 LOCUS
 DEFINITION Homo sapiens clone RP11-127P14, WORKING DRAFT SEQUENCE, 28 unordered pieces.
 ACCESSION AC010770
 VERSION AC010770.4 GI:8096907
 KEYWORDS HTG; HTGS-PHASE1; HTGS-DRAFT.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome, clone RP11-127P14
 Unpublished
 2 (bases 1 to 173645)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,J., Barna,N., Beckerly,K., Boguslavsky,L., Bouckgalter,B., Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArillano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M., Ferrelir,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gage,M., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karacas,A., Klein,J., Lehotzky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Tirrell,A., Vassiliiev,H., Vo,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (22-SEP-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On May 26, 2000 this sequence version replaced gi:6479114.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

1	1166:	contig of 1166 bp in length
*	1167	1266: gap of 100 bp
*	1267	2363: contig of 1097 bp in length
*	2364	2463: gap of 100 bp
*	2464	3608: contig of 1145 bp in length
*	3609	3708: gap of 100 bp
*	3709	5397: contig of 1689 bp in length
*	5398	5497: gap of 100 bp
*	5498	6933: contig of 1436 bp in length
*	6934	7033: gap of 100 bp
*	7034	10416: contig of 3383 bp in length
*	10417	10516: gap of 100 bp
*	10517	14071: contig of 3555 bp in length
*	14072	14171: gap of 100 bp
*	14172	16892: contig of 2721 bp in length
*	16893	16992: gap of 100 bp
*	16993	20502: contig of 3510 bp in length
*	20503	20602: gap of 100 bp
*	20603	24431: contig of 3829 bp in length
*	24432	24531: gap of 100 bp
*	24532	28247: contig of 3716 bp in length
*	28248	28347: gap of 100 bp
*	28348	32668: contig of 4521 bp in length
*	32669	32966: gap of 100 bp
*	32969	36983: contig of 4015 bp in length
*	36984	37083: gap of 100 bp
*	37084	39231: contig of 2148 bp in length
*	39232	39331: gap of 100 bp
*	39332	44516: contig of 5185 bp in length
*	44517	44616: gap of 100 bp
*	44617	48819: contig of 4203 bp in length
*	48820	48919: gap of 100 bp
*	48920	53656: contig of 4737 bp in length
*	53657	53756: gap of 100 bp
*	53757	59428: contig of 5672 bp in length
*	59429	59528: gap of 100 bp
*	59529	65695: contig of 6167 bp in length
*	65696	65795: gap of 100 bp
*	65796	71900: contig of 6105 bp in length
*	71901	72000: gap of 100 bp
*	72001	79394: contig of 7394 bp in length
*	79395	79494: gap of 100 bp
*	79495	90956: contig of 11462 bp in length
*	90957	91056: gap of 100 bp
*	91057	100913: contig of 9857 bp in length

142271. .157726

```

misc_feature      /note="assembly_fragment"
157827..173645
BASE COUNT      52113 a 33977 c 34392 g 50438 t 2725 others
ORIGIN
Query Match      88.0%; Score 22; DB 2; Length 173645;
Best Local Similarity 100.0%; Pred. No. 0.00076;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGATAAG 22
|||||
Db 27538 GGTTCAGTGCAGCCGAGATAAG 27559

RESULT 13
AC068460
LOCUS
DEFINITION      Homo sapiens chromosome 4 clone RP11-70N2 map 4, WORKING DRAFT
ACCESSION      AC068460
VERSION      AC068460.2 GI:8247907
KEYWORDS      HTG; HTGS-PHASE1; HTGS-DRAFT.
SOURCE      Homo sapiens.
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 180315)
AUTHORS      Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE      Homo sapiens chromosome 4, clone RP11-70N2
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 180315)
AUTHORS      Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bida,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campiano,A., Castle,A., Choquel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeAtelliano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Glend,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lenoczky,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severly,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (02-MAY-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 4, 2000 this sequence version replaced gi:7677942.
All repeats were identified using RepeatMasker:
Spit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-----
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: L9488
Center clone name: 70_N_2
-----
Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 174907 bases at least Q40

```

```

Consensus quality: 177492 bases at least Q30
Consensus quality: 178736 bases at least Q20
Insert size: 182000; agarose-fp
Insert size: 179615; sum-of-ontigs
Quality coverage: 4.9 in Q20 bases; agarose-fp
Quality coverage: 4.9 in Q20 bases; sum-of-ontigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1047: contig of 1047 bp in length
* 1048 1147: gap of 100 bp
* 1148 5605: contig of 4458 bp in length
* 5606 5705: gap of 100 bp
* 5706 14016: contig of 8311 bp in length
* 14017 14116: gap of 100 bp
* 14117 21099: contig of 6983 bp in length
* 21100 21199: gap of 100 bp
* 21200 35211: contig of 14012 bp in length
* 35212 35311: gap of 100 bp
* 35312 61645: contig of 26334 bp in length
* 61646 61745: gap of 100 bp
* 61746 108461: contig of 46716 bp in length
* 108462 108561: gap of 100 bp
* 108562 180315: contig of 71754 bp in length.
FEATURES
Location/Qualifiers
Source
1..180315
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-70N2"
/clone_lib="RPC11-11 Human Male BAC"
1..1047
/note="assembly_fragment"
1148..5605
/note="assembly_fragment"
5706..14016
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
14117..21099
/note="assembly_fragment"
clone_end:17
vector_side:right"
21200..35211
/note="assembly_fragment"
35312..61645
/note="assembly_fragment"
61746..108461
/note="assembly_fragment"
108562..180315
/note="assembly_fragment"
BASE COUNT 57120 a 31493 c 31767 g 59234 t
ORIGIN
Query Match 88.0%; Score 22; DB 2; Length 180315;
Best Local Similarity 100.0%; Pred. No. 0.00075;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTTCAGTGCAGCCGAGATAAG 22
|||||
Db 45092 GGTTCAGTGCAGCCGAGATAAG 45113
RESULT 14
AC012100
LOCUS
AC012100 195068 bp DNA linear PRI 21-NOV-2001

```

```

DEFINITION Homo sapiens chromosome 15 clone RP11-507J18 map 15q21.1, complete
sequence.
ACCESSION AC012100
VERSION AC012100.7 GI:17027246
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 195068)
AUTHORS Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
Pate,D. and Hood,L.
TITLE Sequencing of human chromosome 15 D15S146-D15S117 region
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 195068)
AUTHORS Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
Bloom,S., Dors,M., Dickhoff,R., Harrison,G., James,R., Madan,A.,
Owen,M.P., Ratcliffe,A., Shaffer,T. and Hood,L.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-1999) Multimegabase Sequencing Center, University
of Washington, PO BOX 357730, Seattle, WA 98195, USA
REFERENCE 3 (bases 1 to 195068)
AUTHORS Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
Pate,D. and Hood,L.
TITLE Direct Submission
JOURNAL Submitted (21-NOV-2001) Multimegabase Sequencing Center, Institute
for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA
98105, USA
COMMENT On Nov 21, 2001 this sequence version replaced gi:13654329.
----- Genome Center
Center: Multimegabase Sequencing Center
Center code: UWMSC
Web site: http://chroma.mbt.washington.edu/msq_www
Contact: leerowensystemsbiology.org
----- Summary Statistics
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 90% of reads
Chemistry: Dye-primer Big Dye; 10% of reads
Assembly program: Phrap; version 0.990399

Note: Data from overlapping BACs AC084756 (drafting center:
WIBR), AC010770 (drafting center: WIBR), AC036168 (drafting center:
WIBR), and AC021752 (drafting center: UWMSC) were added for
finishing
-----
FEATURES
source Location/Qualifiers
misc_feature 1..195068
misc_feature 1..33375
unsure 42771..42789
unsure 47633..47635
unsure 102471..102472
misc_feature 118026..195068
misc_feature 152001..152888

Homo sapiens chromosome 15 clone RP11-507J18 map 15q21.1, complete
sequence.
ACCESSION AC012100
VERSION AC012100.7 GI:17027246
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 195068)
AUTHORS Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
Pate,D. and Hood,L.
TITLE Sequencing of human chromosome 15 D15S146-D15S117 region
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 195068)
AUTHORS Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
Bloom,S., Dors,M., Dickhoff,R., Harrison,G., James,R., Madan,A.,
Owen,M.P., Ratcliffe,A., Shaffer,T. and Hood,L.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-1999) Multimegabase Sequencing Center, University
of Washington, PO BOX 357730, Seattle, WA 98195, USA
REFERENCE 3 (bases 1 to 195068)
AUTHORS Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
Pate,D. and Hood,L.
TITLE Direct Submission
JOURNAL Submitted (21-NOV-2001) Multimegabase Sequencing Center, Institute
for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA
98105, USA
COMMENT On Nov 21, 2001 this sequence version replaced gi:13654329.
----- Genome Center
Center: Multimegabase Sequencing Center
Center code: UWMSC
Web site: http://chroma.mbt.washington.edu/msq_www
Contact: leerowensystemsbiology.org
----- Summary Statistics
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 90% of reads
Chemistry: Dye-primer Big Dye; 10% of reads
Assembly program: Phrap; version 0.990399

Note: Data from overlapping BACs AC084756 (drafting center:
WIBR), AC010770 (drafting center: WIBR), AC036168 (drafting center:
WIBR), and AC021752 (drafting center: UWMSC) were added for
finishing
-----
FEATURES
source Location/Qualifiers
misc_feature 1..195068
misc_feature 1..33375
unsure 42771..42789
unsure 47633..47635
unsure 102471..102472
misc_feature 118026..195068
misc_feature 152001..152888

```

```

misc_feature 165872..195068
unsure /note="overlap with CTD-2378E12 AC021752"
193866..193871
BASE COUNT 58212 a 41008 c 40204 g 55644 t
ORIGIN
Query Match 88.0%; Score 22; DB 9; Length 195068;
Best Local Similarity 100.0%; Pred. No. 0.00075;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTTCAGTGGCCGAGATAG 22
|||||
Db 106127 GGTTCAGTGGCCGAGATAG 106148
|||||

RESULT 15
AL355499/c
LOCUS AL355499
DEFINITION Human DNA sequence from clone RP11-328K6 on chromosome 6, complete
sequence.
ACCESSION AL355499.15 GI:10086115
VERSION AL355499
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 198141)
AUTHORS Beasley,O.
TITLE Direct Submission
JOURNAL Submitted (30-NOV-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
COMMENT On Sep 11, 2000 this sequence version replaced gi:10039827.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated repeat sequence elements, where the sequence is
ambiguous, there is an annotation using the 'unsure' feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; SW., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
RP11-328K6 is from the library RPCI-11.2 constructed at the Roswell
Park Cancer Institute by the group of Pieter de Jong. For further
details see http://bacpac.med.buffalo.edu/
VECTOR: pBACE3.6
This sequence is the entire insert of clone RP11-328K6 The true
left end of clone RP11-203H2 is at 153730 in this sequence.
FEATURES
source Location/Qualifiers
1..198141
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-328K6"
/cclone_lib="RPCI-11.2"
53..404
repeat_region
/note="MT1A1 repeat: matches 1..365 of consensus"
repeat_region 547..644
/note="TIGER2 repeat: matches 2398..2500 of consensus"
repeat_region 1036..1337

```

```
repeat_region /note="Alusg repeat: matches 1. .303 of consensus"
1971. .2264
repeat_region /note="Alusx repeat: matches 1. .293 of consensus"
2482. .2708
repeat_region /note="Alusx repeat: matches 2. .225 of consensus"
2847. .3275
repeat_region /note="Trigger2a repeat: matches 1. .433 of consensus"
3385. .3695
repeat_region /note="Alusx repeat: matches 1. .309 of consensus"
4027. .4135
repeat_region /note="L2 repeat: matches 2563. .2685 of consensus"
5269. .5327
repeat_region /note="L2 repeat: matches 2364. .2419 of consensus"
6287. .6583
repeat_region /note="Alusg repeat: matches 1. .297 of consensus"
6836. .6940
repeat_region /note="LIME3A repeat: matches 5996. .6103 of consensus"
7103. .7244
repeat_region /note="MIR repeat: matches 20. .192 of consensus"
7359. .7629
repeat_region /note="MER7A repeat: matches 5. .268 of consensus"
7631. .7918
repeat_region /note="AlusP repeat: matches 1. .288 of consensus"
7919. .7958
repeat_region /note="20 copies 2 mer ca 100% conserved"
8561. .9002
repeat_region /note="MER83 repeat: matches 1. .448 of consensus"
9110. .9410
repeat_region /note="Alusx repeat: matches 1. .301 of consensus"
9411. .9448
repeat_region /note="19 copies 2 mer ta 100% conserved"
9604. .9689
repeat_region /note="MIR repeat: matches 41. .126 of consensus"
9988. .10197
misc_feature /note="L2 repeat: matches 2472. .2671 of consensus"
10438. .10865
/note="Cpg island"
/evidence=not_experimental
repeat_region 12263. .13439
repeat_region /note="L1MC5 repeat: matches 6437. .7589 of consensus"
13548. .13696
/note="MER5B repeat: matches 2. .159 of consensus"
14380. .14549
repeat_region /note="L2 repeat: matches 2534. .2690 of consensus"
14550. .14840
repeat_region /note="Alusg1 repeat: matches 1. .288 of consensus"
14841. .15039
repeat_region /note="L2 repeat: matches 2305. .2534 of consensus"
15038. .15448
repeat_region /note="MT2CB repeat: matches 3. .411 of consensus"
15449. .15807
repeat_region /note="THE1B repeat: matches 1. .364 of consensus"
15808. .15914
repeat_region /note="MT2CB repeat: matches 362. .466 of consensus"
15927. .16206
repeat_region /note="Alusg repeat: matches 1. .278 of consensus"
16208. .16494
repeat_region /note="L2 repeat: matches 2129. .2419 of consensus"
16587. .16715
repeat_region /note="L2 repeat: matches 2575. .2709 of consensus"
17070. .17215
repeat_region /note="MIR repeat: matches 100. .252 of consensus"
17925. .17969
repeat_region /note="L2 repeat: matches 2705. .2750 of consensus"
18038. .18356
repeat_region /note="Alusx repeat: matches 1. .312 of consensus"
18394. .18566
repeat_region /note="FRAM repeat: matches -1. .171 of consensus"
18570. .18625
repeat_region /note="28 copies 2 mer aa 75% conserved"
19223. .19310
repeat_region /note="MT1J repeat: matches 103. .193 of consensus"
19877. .20039
/note="LTR33 repeat: matches 350. .519 of consensus"
20403. .20721
/note="Alusx repeat: matches 1. .308 of consensus"
21963. .22029
/note="L1M4 repeat: matches 5721. .5794 of consensus"
22047. .22340
/note="Aluy repeat: matches 1. .294 of consensus"
22864. .23282
/note="L2 repeat: matches 2027. .2491 of consensus"
23285. .23570
/note="MT1J repeat: matches 211. .514 of consensus"
23748. .24062
/note="MT1J repeat: matches 36. .389 of consensus"
24289. .24401
/note="MER30 repeat: matches 11. .130 of consensus"
24402. .24712
/note="Aluy repeat: matches 1. .307 of consensus"
24713. .24814
/note="MER30 repeat: matches 130. .229 of consensus"
25096. .25391
/note="Alusx repeat: matches 1. .294 of consensus"
25554. .25849
/note="Aluwb repeat: matches 1. .295 of consensus"
26135. .26440
/note="Alusg repeat: matches 1. .312 of consensus"
27172. .27466
/note="Alusx repeat: matches 1. .296 of consensus"
27883. .27997
/note="L1M9 repeat: matches 6171. .6302 of consensus"
28577. .28610
/note="17 copies 2 mer tt 85% conserved"
28778. .28863
/note="MT1-INTERNAL repeat: matches 431. .525 of consensus"
28864. .29250
/note="MSTA repeat: matches 1. .426 of consensus"
29251. .29276
/note="MT1-INTERNAL repeat: matches 404. .431 of consensus"
29277. .29575
/note="AluJb repeat: matches 12. .308 of consensus"
29576. .29620
/note="MT1-INTERNAL repeat: matches 356. .404 of consensus"
29621. .29980
/note="MSTB repeat: matches 1. .357 of consensus"
29981. .30274
/note="Alusx repeat: matches 1. .291 of consensus"
30275. .30344
/note="MSTB repeat: matches 357. .426 of consensus"
30345. .30498
/note="MT1-INTERNAL repeat: matches 246. .356 of consensus"
30601. .30889
/note="Alusg repeat: matches 1. .293 of consensus"
32002. .32290
/note="AlusP repeat: matches 1. .290 of consensus"
32307. .32556
/note="L1ME3A repeat: matches 5894. .6163 of consensus"
32575. .32979
/note="L1ME2 repeat: matches 5754. .6315 of consensus"
32980. .33039
/note="L1ME3A repeat: matches 5849. .5908 of consensus"
33707. .34022
/note="AluJo repeat: matches 2. .312 of consensus"
34185. .34946
/note="L1PAL6 repeat: matches 5368. .6145 of consensus"
35004. .35121
/note="FLAM_A repeat: matches 6. .123 of consensus"
35316. .35620
/note="AluJb repeat: matches 1. .306 of consensus"
36193. .36494
```

```

Query Match      88.0%; Score 22; DB 9; Length 198141;
Best Local Similarity 100.0%; Pred. No. 0.00074;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATAAG 22
|||||
Db 112303 GGTTCAGTGAGCCGAGATAAG 112282

RESULT 16
AL512373/c
LOCUS      AL512373      209790 bp      DNA      linear      HTG 12-SEP-2001
DEFINITION Homo sapiens chromosome 6 clone RP11-564014, *** SEQUENCING IN
PROGRESS ***, in ordered pieces.
ACCESSION  AL512373
VERSION     AL512373.12 GI:15617280
KEYWORDS   HTG; HTGS_PHASE2; HTGS_CANCELLED.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            Smith, M.
            Direct Submission
            Submitted (11-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
            requests: clonerequest@sanger.ac.uk
            On Sep 13, 2001 this sequence version replaced gi:13273887.
            ----- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            ----- Project Information
            Center project name: BA564014
            ----- Summary Statistics
            Sequencing program: XGAP4; version 4.5
            Sequencing vector: plasmid; L08752; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Consensus quality: 209055 bases at least Q40
            Consensus quality: 209344 bases at least Q30
            Consensus quality: 209466 bases at least Q20
            Insert size: 209790; sum-of-contigs
            Insert size: 194715; 2.8% error; agarose-fp
            Quality coverage: 8.31x in Q20 bases; sum-of-contigs Quality
            coverage: 8.96x in Q20 bases; agarose-fp
            -----
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            Location/Qualifiers
            source          1..209790
                           /organism="Homo sapiens"
                           /db_xref="taxon:9606"
                           /chromosome="6"
                           /clone="RP11-564014"
                           /clone_lib="RPC1-11.2"
            misc_feature   1..209790
                           /note="assembly_fragment:04568"
BASE COUNT  58386 a 41949 c 43662 g 65793 t
ORIGIN
Query Match      88.0%; Score 22; DB 2; Length 209790;
Best Local Similarity 100.0%; Pred. No. 0.00074;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATAAG 22
|||||
Db 77902 GGTTCAGTGAGCCGAGATAAG 77881

RESULT 17

```

```

AX070981/c
LOCUS      AX070981      355 bp      DNA      linear      PAT 25-JAN-2001
DEFINITION Sequence 1453 from Patent WO0102568.
ACCESSION  AX070981
VERSION     AX070981.1 GI:12581254
KEYWORDS
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            Williams, L.T., Escobedo, J., Innis, M.A., Garcia, P.D., Klinger, J.,
            Kassam, A., Reinhard, C., Randazzo, F., Kennedy, G.C., Pot, D.,
            Lamson, G., Drmanac, R., Crkenjakov, R., Drmanac, S., Dickson, M.,
            Labat, I., Leshkowitz, D., Kita, D., Garcia, V. and Strache-Crain, B.
            Human genes and gene expression products
            Patent: WO 0102568-A 1453 11-JAN-2001;
            CHIRON CORPORATION (US) ; HYSEQ, INC. (US)
            Location/Qualifiers
            source          1..355
                           /organism="Homo sapiens"
                           /db_xref="taxon:9606"
BASE COUNT  87 a 94 c 72 g 102 t
ORIGIN
Query Match      84.0%; Score 21; DB 6; Length 355;
Best Local Similarity 100.0%; Pred. No. 0.0066;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATAA 21
|||||
Db 60 GGTTCAGTGAGCCGAGATAA 40

RESULT 18
HSPA27G11/c
LOCUS      HSPA27G11      520 bp      DNA      linear      PRI 23-AUG-1996
DEFINITION H.sapiens flow-sorted chromosome 6 HindIII fragment, SC6PA27G11.
ACCESSION  Z79279
VERSION     Z79279.1 GI:1508557
KEYWORDS   Anonymous marker, single read.
SOURCE     Homo sapiens.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            Mungall, A.J., Huckle, E., Langford, C., Ross, M.T. and Rice, C.M.
            Direct Submission
            Submitted (22-AUG-1996) The Sanger Centre, Wellcome Trust Genome
            Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
            humquery@sanger.ac.uk
            Vector: pBSIISK+.
            Location/Qualifiers
            source          1..520
                           /organism="Homo sapiens"
                           /db_xref="taxon:9606"
                           /chromosome="6"
                           /clone="SC6PA27G11"
                           /sex="female"
                           /tissue_type="EBV lymphoblastoid cell line"
                           /clone_lib="SC6PA"
                           /dev_stage="adult"
                           /note="The estimated purity of the flow-sorted chromosome
                           6 library is >97%"
BASE COUNT  101 a 137 c 85 g 197 t
ORIGIN
Query Match      84.0%; Score 21; DB 9; Length 520;
Best Local Similarity 100.0%; Pred. No. 0.0063;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATAA 21
|||||

```

```

Db 419 GGTTCAGTGCAGCCGAGATAA 399

RESULT 19
LOCUS AX195315/c
DEFINITION Sequence 5308 from Patent WO0214500.
ACCESSION AX195315
VERSION AX195315
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Escobedo,J., Garcia,P.D., Sudduth-Klinger,J., Reinhard,C.,
Randazzo,F., Lamson,G., Scott,E.M., Zhang,G., Kassam,A., Pot,D. and
Labat,I.
TITLE Human genes and gene expression products
JOURNAL Patent: WO 0214500-A 5308 21-FEB-2002;
MEDLINE CHIRON CORPORATION (US) ; Hyseq Inc. (US)
PUBMED Location/Qualifiers
1..608
FEATURES
source
BASE COUNT 153 a 126 c 114 g 215 t
ORIGIN
Query Match 84.0%; Score 21; DB 6; Length 608;
Best Local Similarity 100.0%; Pred. No. 0.0062;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGCAGCCGAGATAA 21
|||||
Db 501 GGTTCAGTGCAGCCGAGATAA 481

RESULT 20
LOCUS AX195315/c
DEFINITION Sequence 19 from Patent WO0151631.
ACCESSION AX195315
VERSION AX195315
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Reske-Kunz,A., Ross,X., Ross,R. and Bros,M.
TITLE Regulatory sequence for the specific expression in dendritic cells
and uses thereof
JOURNAL Patent: WO 0151631-A 19 19-JUL-2001;
MEDLINE Reske-Kunz, Angelika (DE) ; Ross, Xiaolan (DE) ; Ross, Ralf (DE) ;
PUBMED Bros, Matthias (DE)
FEATURES
source
BASE COUNT 230 a 330 c 387 g 295 t
ORIGIN
Query Match 84.0%; Score 21; DB 6; Length 1242;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGCAGCCGAGATAA 21
|||||
Db 522 GGTTCAGTGCAGCCGAGATAA 502

RESULT 21
LOCUS HUMER2B/c
DEFINITION Human tyrosine kinase-type receptor (HER2) gene, partial cds.
ACCESSION M12036
VERSION M12036.1 GI:183988
KEYWORDS tyrosine kinase.
SOURCE Human fetal DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Coussens,L., Yang-Feng,T.L., Liao,Y.-C., Chen,E., Gray,A.,
McGrath,J., Seeburg,P.H., Libermann,T.A., Schlessinger,J.,
Francke,U., Levinson,A. and Ullrich,A.
TITLE Tyrosine Kinase receptor with extensive homology to EGF receptor
shares chromosomal location with neu oncogene
JOURNAL Science 230 (4730), 1132-1139 (1985)
MEDLINE 86070181
PUBMED 2999974
FEATURES
source
Location/Qualifiers
1..1838
/organism="Homo sapiens"
/db_xref="taxon:9606"
join(1..92,806..991,1129..1284,1405..1488)
/partial
/note="HER2 receptor (AA at 3)"
/codon_start=3
/protein_id="AAA35978.1"
/db_xref="GI:183988"
/translation="IPDGENVKIPVAIKVLRNTPSKANKEILDEAYVMAGVSPYVS
RLIGTCTSTVOLVTLMPYGCGLLDHVRNKGRLGSDLLNWCQIAKMGVLEDEVL
VHRLAARNVLYKSPNHVKITDFGLARLLIDIDETEHADGSGKVPFKWMALESILRRF
THQSDWNSYGV"
exon <1..92
exon 1..92
/note="exon x"
1..92
/partial
/note="exon x"
93..805
/note="intron A"
<806..991
/note="exon x+1"
806..991
/note="exon x+1"
992..1128
/note="intron B"
<1129..1284
/note="exon x+2"
1285..1404
/note="intron C (no splice consensus); putative"
1405..>1485
/note="exon x+3"
1486..>1838
/note="intron D (no splice consensus); putative"
BASE COUNT 358 a 487 c 551 g 442 t
ORIGIN Chromosome 17q21-q22.

Query Match 84.0%; Score 21; DB 9; Length 1838;
Best Local Similarity 100.0%; Pred. No. 0.0055;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGCAGCCGAGATAA 21
|||||
Db 282 GGTTCAGTGCAGCCGAGATAA 262

RESULT 22
LOCUS AX195330/c
DEFINITION Sequence 34 from Patent WO0151631.
ACCESSION AX195330
VERSION AX195330
KEYWORDS
SOURCE human.

```

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Reske-Kunz,A., Ross,X., Ross,R. and Bros.M.
TITLE Regulatory sequence for the specific expression in dendritic cells
and uses thereof
JOURNAL Patent: WO 0151631-A 34 19-JUL-2001;
Reske-Kunz, Angelika (DE) ; Ross, Xiaolan (DE) ; Ross, Ralf (DE) ;
Bros, Matthias (DE)
FEATURES
Source Location/Qualifiers
1..2843
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 499 a 909 c 827 g 608 t
ORIGIN
Query Match 84.0%; Score 21; DB 6; Length 2843;
Best Local Similarity 100.0%; Pred. No. 0.0053;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTTCAGTGGAGCCGAGATAA 21
|||||
Db 1607 GGTTCAGTGGAGCCGAGATAA 1587
|||||

RESULT 23
HSAC000981/c HSAC000981 3761 bp DNA linear PRI 09-APR-1997
LOCUS Homo sapiens (subclone 2_c6 from p1 H31) DNA sequence, complete
ACCESSION AC000981
VERSION AC000981.1 GI:1930997
KEYWORDS HTG.
SOURCE Homo sapiens (Subclones in pOT2 from p1 clone H31) DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Martin,C.H., Arcaina,T.T., Bondoc,M.M., Chiang,A., Critz,P.A.,
Davis,C.A., Doyle,C.M., Ericsson,C.L., Farfan,D.E., Gunning,K.M.,
Houston,K.A., Jaklevic,M.A., Kadner,K.E., Kim,K., Kim,S.F.,
Ko,C.L., Lewis,K.D., Li,M., Lindquist,K.J., Lomotan,M.A.,
Lustre,V.M., Machrus,M.U., Mayeda,C.A., Miguel,T.M., Miller,C.A.,
Mok,M.S., Pacleb,J.M., Patel,S.G., Santos,R.F., Subramanian,S.,
Wan,K.H., Whitelaw,K.R., Yee,A., Yen,R.T., Yu,C. and Palazzolo,M.J.
Sequencing of human chromosome 5q
Unpublished (1996)
2 (bases 1 to 3761)
Martin,C.H., Arcaina,T.T., Bondoc,M.M., Chiang,A., Critz,P.A.,
Davis,C.A., Doyle,C.M., Ericsson,C.L., Farfan,D.E., Gunning,K.M.,
Houston,K.A., Jaklevic,M.A., Kadner,K.E., Kim,K., Kim,S.F.,
Ko,C.L., Lewis,K.D., Li,M., Lindquist,K.J., Lomotan,M.A.,
Lustre,V.M., Machrus,M.U., Mayeda,C.A., Miguel,T.M., Miller,C.A.,
Mok,M.S., Pacleb,J.M., Patel,S.G., Santos,R.F., Subramanian,S.,
Wan,K.H., Whitelaw,K.R., Yee,A., Yen,R.T., Yu,C. and Palazzolo,M.J.
Submitted (09-APR-1997)
Sequence submitted by:
Human Genome Center
Lawrence Berkeley National Laboratory, MS 74-157
Berkeley, CA 94720
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Website (http://www.hgc.lbl.gov/sequence-archive.html) or
send email to humangenome.lbl.gov.
Location/Qualifiers
1..3761
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="p1 H31 (5219)"
/sub_clone="2_c6"
FEATURES
Source

```

```

BASE COUNT 1040 a 824 c 639 g 1258 t
ORIGIN
Query Match 84.0%; Score 21; DB 9; Length 3761;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTTCAGTGGAGCCGAGATAA 21
|||||
Db 978 GGTTCAGTGGAGCCGAGATAA 958
|||||

RESULT 24
AX195368/c AX195368 16951 bp DNA linear PAT 28-AUG-2001
LOCUS Sequence 72 from Patent WO0151631.
ACCESSION AX195368
VERSION AX195368.1 GI:15385917
KEYWORDS human.
SOURCE Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Reske-Kunz,A., Ross,X., Ross,R. and Bros.M.
TITLE Regulatory sequence for the specific expression in dendritic cells
and uses thereof
JOURNAL Patent: WO 0151631-A 72 19-JUL-2001;
Reske-Kunz, Angelika (DE) ; Ross, Xiaolan (DE) ; Ross, Ralf (DE) ;
Bros, Matthias (DE)
FEATURES
Source Location/Qualifiers
1..16951
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 3224 a 4928 c 5200 g 3599 t
ORIGIN
Query Match 84.0%; Score 21; DB 6; Length 16951;
Best Local Similarity 100.0%; Pred. No. 0.0043;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTTCAGTGGAGCCGAGATAA 21
|||||
Db 14694 GGTTCAGTGGAGCCGAGATAA 14674
|||||

RESULT 25
AY044229/c AY044229 16952 bp DNA linear PRI 15-SEP-2001
LOCUS Homo sapiens fascic gene, complete cds.
ACCESSION AY044229
VERSION AY044229.1 GI:15625240
KEYWORDS Homo sapiens.
SOURCE Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Bros,M., Ross,X.L., Reske-Kunz,A.B. and Ross,R.
TITLE Human Fascin Gene Sequence
JOURNAL Unpublished
2 (bases 1 to 16952)
Bros,M., Ross,X.L., Reske-Kunz,A.B. and Ross,R.
Direct Submission
Submitted (08-JUL-2001) Clinical Research Unit, Department of
Dermatology, Johannes Gutenberg University, Obere Zahlbacher Str.
63, Mainz 55131, Germany
Location/Qualifiers
1..16952
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7p22"
FEATURES
Source

```


TATA_signal
mRNA

/clone="RPCIP704C24766Q3/4"
2924. .2930

Join(2958. .3910,13399. .13555,13638. .13759,14005. .14172,
15415. .16791)

/product="fascin"

Join(3079. .3910,13399. .13555,13638. .13759,14005. .14172,
15415. .15617)

/note="actin bundling protein"

/codon_start=1

/product="fascin"

/protein_id="AA01526.1"

/db_xref="GI:15625241"

/translation="MTANGTAEAVIQFGLINGNKLYTAEAFKVNASASSLKXKO

ITLQEPDDEAGSAACLRSLHGLRYLAADKGNVTCREVPDPCDFLIIVADGGRWS

LQSAHRRYEGGEDRLSCFAQTVSPAERKSVHIAHPQNTVTKRYIAHLGARA

DEIADVDPMGVDLSILTAPOQSYSTVADHRLDGLRLVAREPATGYTLFFRS

GKFAFDCEGRVLASGSGELKAGKATYKDELFALQSCAQVYLOAANERNVSR

QGMDSANODEITDQTEFLEIDRDKKCAFTHTGKYWTLTATGGVQSTASSKASC

YFLEMDRRLITLRASNGKFVTKKNGLAASVETAGDSFLFKMLINRPIIVRFGH

GFICRKVTGLDANRSYDVFQLEFNDGAYNKDSGKYWTYVSDSAVITSSGDTFVD

FFFECDYNKVAIKVGGRYLKGHAGLVKASNFVDFASLWEY"

BASE COUNT 3224 a 4928 c 5201 g 3599 t

ORIGIN

Query Match

Best Local Similarity 84.0%; Score 21; DB 9; Length 16952;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGAGCGGAGATAA 21

|||||

Db 14694 GGTTCAGTGAGCGGAGATAA 14674

RESULT 26

AC092149/c

LOCUS

DEFINITION Homo sapiens chromosome RPCI-11 clone RP11-96J6, WORKING DRAFT

SEQUENCE, 18 unordered pieces.

AC092149

VERSION AC092149.1 GI:14523004

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 24144)

Waterston,R.H.

Unpublished

2 (bases 1 to 24144)

Waterston,R.H.

Direct Submission

Submitted (22-JUN-2001) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

COMMENT

----- Genome Center -----

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc/index.shtml>

----- Project Information -----

Center project name: H_NH0096J06

----- Summary Statistics -----

Sequencing vector: M13; 0%

Sequencing vector: plasmid; 100%

Chemistry: Dye-primer ET; 0% of reads

Assembly: Dye-terminator Big Dye; 100% of reads

Consensus quality: 19352 bases at least Q40

Consensus quality: 20525 bases at least Q30

Consensus quality: 21376 bases at least Q20

Insert size: 194000; agarose-fp

Insert size: 22444; sum-of-contigs
Quality coverage: 0.17 in Q20 bases; agarose-fp
Quality coverage: 1.55 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1181: contig of 1181 bp in length

1182 1281: gap of unknown length

1282 2421: contig of 1140 bp in length

2422 2521: gap of unknown length

2522 3771: contig of 1250 bp in length

3772 3871: gap of unknown length

3872 5067: contig of 1196 bp in length

5068 5167: gap of unknown length

5168 6341: contig of 1174 bp in length

6342 6441: gap of unknown length

6442 7566: contig of 1125 bp in length

7567 7666: gap of unknown length

7667 8713: contig of 1047 bp in length

8714 8813: gap of unknown length

8814 9929: contig of 1116 bp in length

9930 10029: gap of unknown length

10030 11336: contig of 1307 bp in length

11337 11436: gap of unknown length

11437 12571: gap of unknown length

12572 12672: contig of 1043 bp in length

12673 13814: gap of unknown length

13815 14889: contig of 1075 bp in length

14890 14989: gap of unknown length

14990 16070: contig of 1081 bp in length

16071 16170: gap of unknown length

16171 17718: contig of 1548 bp in length

17719 19379: contig of 1561 bp in length

19380 19479: gap of unknown length

19480 20568: contig of 1089 bp in length

20569 20668: gap of unknown length

20670 22125: contig of 1457 bp in length

22126 22225: gap of unknown length

22226 24144: contig of 1919 bp in length.

Location/Qualifiers

1. .24144

/organism="Homo sapiens"

/db_xref="Taxon:9606"

/chromosome="RPCI-11"

/clone="RP11-96J6"

1. .1181

/note="assembly_name:Contig19"

1282. .2421

/note="assembly_name:Contig22"

2522. .3771

/note="assembly_name:Contig24"

3872. .5067

/note="assembly_name:Contig33"

5168. .6341

/note="assembly_name:Contig35"

6442. .7566

/note="assembly_name:Contig36"

7667. .8713

/note="assembly_name:Contig38"

8814. .9929

/note="assembly_name:Contig40"

10030. .11336

/note="assembly_name:Contig41"

11437. .12571

/note="assembly_name:Contig42"

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

```

misc_feature 12672..13714
/note="assembly_name:Contig43"
misc_feature 13815..14889
/note="assembly_name:Contig44"
misc_feature 14990..16070
/note="assembly_name:Contig46"
misc_feature 16171..17718
/note="assembly_name:Contig47"
misc_feature 17819..19379
/note="assembly_name:Contig48"
misc_feature 19480..20568
/note="assembly_name:Contig51"
misc_feature 20669..22125
/note="assembly_name:Contig52"
misc_feature 22226..24144
/note="assembly_name:Contig53"
BASE COUNT 6285 a 4871 c 4766 g 6518 t 1704 others
ORIGIN

Query Match 84.0%; Score 21; DB 2; Length 24144;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATAA 21
|||||
Db 14451 GGTTCAGTGAGCCGAGATAA 14431

RESULT 27
HSL112A12 25889 bp DNA linear PRI 23-NOV-1999
LOCUS
DEFINITION Human DNA sequence from cosmid L112A12, Huntington's Disease.
ACCESSION Z68225
VERSION Z68225.1 GI:1122886
KEYWORDS 4p16.3; CpG island.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
I (bases 1 to 25889)
Direct Submission
Submitted (12-DEC-1995) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RO, England. E-mail enquiries: humquery@sanger.ac.uk
IMPORTANT:
This sequence is the entire insert of clone L112A12. L112A12 is
from cosmid library LA04NC01 constructed at the Human Genome
Center, Los Alamos National Laboratory, NM 87545 under the auspices
of the U.S. Department of Energy. The library was constructed using
flow-sorted human chromosome 4 from a Hamster-Human hybrid cell
line (UV20HL21-27) containing human chromosomes 4, 8 and 21.
VECTOR: scos1
L112A12 is contained in a clone contig spanning
2Mb which is described in Baxendale et al, Nature Genetics 4 ( 1993
) 181- 186. See also Myers et al, Cytogenet Cell Genet. 66 ( 1994
) 218-230.
FEATURES
source
Location/Qualifiers
1..25889
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4p16.3"
/clone="LA04NC01-L112A12"
/cell_line="UV20HL21-27"
/cell_lib="LA04NC01"
133..200
/note="2 copies of 34 mer 90 % conserved"
repeat_region 167..210
/note="11 copies of 4 mer 86 % conserved"
repeat_region 1102..1169
/note="2 copies of 34 mer 97 % conserved; 34 copies of 2
mer 97 % conserved"

```

```

repeat_region 1104..1167
/note="16 copies of 4 mer 98 % conserved"
repeat_region 1829..1872
/note="11 copies of 4 mer 100 % conserved"
repeat_region 3065..3170
/note="2 copies of 53 mer 93 % conserved"
repeat_region 4921..5272
/note="MLTIC element fragment"
5142..5258
/note="MLTIC element fragment"
5775..5818
/note="11 copies of 4 mer 89 % conserved"
repeat_region 7952..7997
/note="MLTIC element fragment"
8593..8698
/note="2 copies of 53 mer 95 % conserved"
repeat_region 9263..9553
/note="Alu repeat: matches 1..308 of consensus"
repeat_region 12998..13096
/note="MIR element fragment"
15090..15264
/note="MER42C element fragment"
15272..15563
/note="Alu repeat: matches 1..308 of consensus"
repeat_region 15588..15879
/note="Alu repeat: matches 1..308 of consensus"
repeat_region 15898..16003
/note="MER42C element fragment"
16381..16530
/note="Li element fragment"
repeat_region 16755..16829
/note="MLTIC element fragment"
repeat_region 17329..17437
/note="MLTIC element fragment"
repeat_region 17331..17504
/note="MLTIC element fragment"
repeat_region 17513..17599
/note="MLTIC element fragment"
repeat_region 19285..19573
/note="Alu repeat: matches 1..308 of consensus"
19863..20421
/note="match: 5' EST T57716 clone 79379; Paired with EST
T57677 matching this cosmid"
misc_feature 21469..21672
/note="match: 3' EST T57677 clone 79379; Paired with EST
T57716 matching this cosmid"
repeat_region 22060..22125
/note="6 copies of 34 mer 94 % conserved"
repeat_region 24247..25887
/note="MIR element fragment"
misc_feature 24247..25887
/note="putative CpG island"
BASE COUNT 5888 a 7239 c 7460 g 5302 t
ORIGIN

Query Match 84.0%; Score 21; DB 9; Length 25889;
Best Local Similarity 100.0%; Pred. No. 0.0041;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATAA 21
|||||
Db 15804 GGTTCAGTGAGCCGAGATAA 15824

RESULT 28
AC108170
LOCUS
DEFINITION Homo sapiens BAC clone RP11-1401p21 from 4, complete sequence.
ACCESSION AC108170
VERSION AC108170.4 GI:19807987
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens

```



```

repeat_region      /rpt_family="L1"
16188..16283
repeat_region      /rpt_family="L2"
16284..16383
repeat_region      /rpt_family="GA-rich"
18047..18222
repeat_region      /rpt_family="MER1_type"
18290..18349
repeat_region      /rpt_family="L2"
18983..19294
repeat_region      /rpt_family="ERV1"
19963..20147
repeat_region      /rpt_family="MIR"
20846..21071
repeat_region      /rpt_family="L1"
21782..21813
repeat_region      /rpt_family="AT-rich"
22397..22473
repeat_region      /rpt_family="CR1"
22475..22522
repeat_region      /rpt_family="L2"
22514..22606
repeat_region      /rpt_family="CR1"
22650..22736
repeat_region      /rpt_family="L2"
22737..23074
repeat_region      /rpt_family="MER1_type"
23075..23127
repeat_region      /rpt_family="L2"
23238..23453
repeat_region      /rpt_family="MIR"
23485..23618
repeat_region      /rpt_family="L2"
25127..25231
repeat_region      /rpt_family="MALR"
25363..25914
repeat_region      /rpt_family="L2"
25892..26046
repeat_region      /rpt_family="L2"
26047..26912
repeat_region      /rpt_family="ERV1"
BASE COUNT 8951 a 4918 c 5017 g 8027 t
ORIGIN
Query Match      84.0%; Score 21; DB 9; Length 26913;
Best Local Similarity 100.0%; Pred. No. 0.0041;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTGGCAGTCAGCCGAGATAA 21
|||||
Db 15582 GGTGGCAGTCAGCCGAGATAA 15602
|||||

RESULT 29
AL354801/c
LOCUS
DEFINITION Human DNA sequence from clone RP4-700G13 on chromosome 20. Contains the 5' part of the BCAS1 gene for breast carcinoma amplified sequence 1, ESTs, STSs and GSSs, complete sequence.
ACCESSION AL354801
VERSION AL354801.5 GI:8546654
KEYWORDS HNG; BCAS1; breast carcinoma.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 31163)
AUTHORS Griffiths, C.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
COMMENT On Jun 14, 2000 this sequence version replaced gi:8250083.

```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr20>

IMPORTANT: This sequence is not the entire insert of clone RP4-700G13. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone RP5-1075G21 is at 31064 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP4-700G13 is from the library RPCI-4 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

FEATURES

```

source
1..31163
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/clone="RP4-700G13"
/clone_lib="RPCI-4"
1..100
misc_feature
/notes="true right end of clone Em.AC004501 is at 100 in this sequence"
2..63
repeat_region
/notes="MTIB repeat: matches 330..390 of consensus"
320..618
repeat_region
/notes="AluX repeat: matches 1..300 of consensus"
1278..1425
repeat_region
/notes="MER31A repeat: matches 1..163 of consensus"
2920..3084
repeat_region
/notes="MIR repeat: matches 10..174 of consensus"
3437..3585
gene
/notes="MIR repeat: matches 35..191 of consensus"
complement(3660..3771)
gene
/genes="BCAS1"
complement(<3660..3771)
mRNA
/genes="BCAS1"
/product="dJ700G13.1 (breast carcinoma amplified sequence 1)"
/notes="continues in Em:AC004501 and Em:AC005220"
match: CDNAS: Em:AF041260
match: ESTs: Em:BF678192
/evidence="not_experimental"
5321..5618
repeat_region
/notes="AluX repeat: matches 5..302 of consensus"
5839..6520
repeat_region
/notes="11 copies 62 mer 96% conserved"
6520..7416
repeat_region
/notes="MER45B repeat: matches 1..931 of consensus"
7417..7906
repeat_region
/notes="Tigger2a repeat: matches 1..434 of consensus"
7907..7960
repeat_region
/notes="MER45B repeat: matches 931..982 of consensus"
8723..8953
repeat_region
/notes="MIR repeat: matches 15..248 of consensus"
8951..9030

```

```
repeat_region /note="MIR repeat: matches 12. .92 of consensus"
9503. .9745
repeat_region /note="MIR repeat: matches 8. .262 of consensus"
9844. .10070
repeat_region /note="AluSc repeat: matches 64. .290 of consensus"
10210. .10307
repeat_region /note="MIR repeat: matches 47. .147 of consensus"
10471. .10679
repeat_region /note="MER20 repeat: matches 1. .218 of consensus"
11021. .11373
repeat_region /note="L1MB4 repeat: matches 5816. .6181 of consensus"
11374. .11441
repeat_region /note="MER53 repeat: matches 105. .179 of consensus"
11379. .11458
repeat_region /note="MER53 repeat: matches 35. .113 of consensus"
11457. .11594
repeat_region /note="MER53 repeat: matches 1. .122 of consensus"
11952. .11997
repeat_region /note="23 copies 2 mer aa 84% conserved"
12155. .12335
repeat_region /note="MER58A repeat: matches 41. .224 of consensus"
12392. .12542
repeat_region /note="MIR repeat: matches 20. .182 of consensus"
12633. .12933
repeat_region /note="AluX repeat: matches 1. .311 of consensus"
12944. .13496
repeat_region /note="Charliel repeat: matches 2217. .2760 of consensus"
14391. .14699
repeat_region /note="AluDo repeat: matches 1. .302 of consensus"
15229. .15423
repeat_region /note="AluDo repeat: matches 2. .308 of consensus"
16634. .16935
repeat_region /note="MIR repeat: matches 39. .146 of consensus"
17464. .17782
repeat_region /note="match: GSS: Em:AL172133"
complement(17467. .17750)
repeat_region /note="match: STS: Em:G05108"
17515. .17704
repeat_region /note="match: GSS: Em:AQ530947"
complement(17551. .17766)
repeat_region /note="match: GSS: Em:AL229216"
17555. .17704
repeat_region /note="match: GSS: Em:AQ530949"
17568. .17791
repeat_region /note="match: STS: Em:AQ029014"
complement(17578. .17778)
repeat_region /note="match: GSS: Em:AL003817"
17580. .17778
repeat_region /note="match: GSS: Em:AL306181 Em:AL348631"
complement(17625. .17823)
repeat_region /note="match: STS: Em:HS1024HT"
17634. .17777
repeat_region /note="match: GSS: Em:AQ078596"
17650. .17808
repeat_region /note="match: GSS: Em:AQ740298"
17793. .18142
repeat_region /note="L2 repeat: matches 2366. .2705 of consensus"
18922. .18986
repeat_region /note="MIR repeat: matches 80. .257 of consensus"
19012. .19028
repeat_region /note="MIR repeat: matches 4. .19 of consensus"
19029. .19156
repeat_region /note="MER85 repeat: matches 1. .139 of consensus"
19157. .19368
repeat_region /note="MIR repeat: matches 19. .235 of consensus"
19722. .19944
repeat_region /note="MIR repeat: matches 25. .250 of consensus"
20583. .20864
repeat_region /note="AluSc repeat: matches 1. .297 of consensus"
```

```
repeat_region 20927. .21114
/note="LTR33 repeat: matches 3. .184 of consensus"
22207. .22331
repeat_region /note="L2 repeat: matches 2572. .2702 of consensus"
22450. .22930
repeat_region /note="LTR32 repeat: matches 17. .471 of consensus"
23085. .23281
repeat_region /note="MLTJ repeat: matches 312. .516 of consensus"
24302. .24589
repeat_region /note="AluCb repeat: matches 1. .301 of consensus"
25088. .25185
repeat_region /note="49 copies 2 mer ct 68% conserved"
25726. .25766
repeat_region /note="MER5B repeat: matches 137. .178 of consensus"
25822. .25876
repeat_region /note="L1MEC repeat: matches 2239. .2291 of consensus"
25877. .26192
repeat_region /note="AluSx repeat: matches 1. .310 of consensus"
26193. .26587
repeat_region /note="L1MEC repeat: matches 1713. .2239 of consensus"
26705. .26758
repeat_region /note="MER5B repeat: matches 122. .174 of consensus"
27852. .27945
repeat_region /note="MIR repeat: matches 101. .197 of consensus"
27946. .28060
repeat_region /note="MER85 repeat: matches 1. .137 of consensus"
28061. .28093
repeat_region /note="MIR repeat: matches 70. .102 of consensus"
28268. .28716
repeat_region /note="match: GSS: Em:AQ320817"
28278. .28716
repeat_region /note="match: GSS: Em:AQ320816"
28302. .28479
repeat_region /note="L1PA5 repeat: matches 5966. .6141 of consensus"
28704. .29014
repeat_region /note="AluSg repeat: matches 1. .310 of consensus"
```

Query Match 84.0%; Score 21; DB 9; Length 31163;
Best Local Similarity 100.0%; Pred.No. 0.004;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGTTCAGTGAGCCGAGATAA 21
|||||
Db 5405 GGTTCAGTGAGCCGAGATAA 5385

RESULT 30

AC005331
LOCUS

DEFINITION

AC005331

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 35581)

Lamerdin, J.E., McCready, P.M., Skowronski, E., Adamson, A.W.,

Burkhardt-Schultz, K., Gordon, L., Kyle, A., Ramirez, M., Stilwagen, S.,

Phan, H., Velasco, N., Do, L., Regalia, W., Terry, A., Barnes, J.,

Danganan, L., Poundstone, P., Christensen, M., Georgescu, A., Avila, J.,

Liu, S., Attix, C., Andreise, T., Frankheim, M., Amico-Keller, G.,

Coelefeld, J., Duarte, S., Lucas, S., Bruce, R., Thomas, P., Quan, G.,

Kronmiller, B., Arellano, A., Olsen, A.S. and Carrano, A.V.

Sequence analysis of a 3.5 Mb contig in human 19p13.3 containing a

serine protease gene cluster

Unpublished

JOURNAL

REFERENCE

AUTHORS

Lamerdin, J.E.

Direct Submission

Submitted (28-JUL-1998) Joint Genome Institute, Lawrence Livermore


```

repeat_region complement(14942..15092)
repeat_region /rpt_family="LIMB7"
repeat_region 15107..15403
repeat_region /rpt_family="AluJo"
repeat_region 15416..15585
repeat_region /rpt_family="AluSg/x"
repeat_region 15604..15904
repeat_region /rpt_family="AluSx"
repeat_region 15936..16098
repeat_region /rpt_family="AluSg/x"
repeat_region complement(16215..16446)
repeat_region /rpt_family="LINE2"
repeat_region complement(16463..16631)
repeat_region /rpt_family="AluJb"
repeat_region complement(16632..16934)
repeat_region /rpt_family="AluSp"
repeat_region complement(16937..17068)
repeat_region /rpt_family="AluJb"
repeat_region complement(17373..17679)
repeat_region /rpt_family="AluSg"
repeat_region complement(17908..17985)
repeat_region /rpt_family="GC-rich"
repeat_region 18004..18063
repeat_region /rpt_family="(CGG)n"
repeat_region complement(18056..18134)
repeat_region /rpt_family="(CGG)n"
repeat_region complement(18182..18237)
repeat_region /rpt_family="GC-rich"
misc_feature complement(18240..18367)

```

```

Query Match 84.0%; Score 21; DB 9; Length 35581;
Best local Similarity 100.0%; Pred. No. 0.004;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGTGAGTGGCGGAGATAA 21
|||||
Db 15820 GGTGAGTGGCGGAGATAA 15840

```

```

RESULT 31
AC006138
LOCUS AC006138 37448 bp DNA linear PRI 05-DEC-1998
DEFINITION Homo sapiens clone UWGC:y54c194 from 6p21, complete sequence.
ACCESSION AC006138
VERSION AC006138.1 GI:3970958
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Janer, M., Guillaudeau, T., Vu, O., Kutayavin, T., Harter, H. and
Geraghty, D.E.
TITLE Large scale sequence analysis of the human MHC class I region
JOURNAL Unpublished (1998)
REMARK Fred Hutchinson Cancer Research Center
The Clinical Research Division
1100 Fairview Ave. N., P.O. Box 19024
Seattle, WA 98109-1024
Contact: Daniel E. Geraghty (geraghty@fhrc.org)
2 (bases 1 to 37448)
Geraghty, D.E. and Olson, M.V.
Direct Submission
Submitted (05-DEC-1998) Human Genome Center, University of
Washington, Box 352145, Seattle, WA 98195, USA
University of Washington Human Genome Center
Box 352145 Seattle, WA 98195
Contact: Daniel E. Geraghty (geraghty@fhrc.org)
COMMENT Overlapping Sequences:
5': UWGC:y54c222 (Genbank Accession: AC006049)
3': UWGC:y55c174
-----
Sequence Quality Assessment:

```

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

Double stranded (DS) coverage: 63.0%
DS or two chemistry coverage: 99.8%
Single stranded regions: 2

Sequence Validation:

This sequence has been validated by Multiple Complete Digest Mapping. Comparison of the experimentally derived map digest fragments with sequence-predicted fragments is given below. Small fragments below a variable cutoff (approximately 400-600bp) are not mapped and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragment groups are separated by dashed lines.

BgIII			HindIII			NsiI		
Map	Seq	Map	Seq	Map	Seq	Map	Seq	Seq
10275.50	9934.00	4611.50	4460.00	2433.00	2383.00			
7579.50	7368.00	6983.50	6709.00	1325.00	1321.00			
2393.50	2344.00	2265.33	2230.00	2932.50	2884.00			
1196.38	1179.00	16560.50	16097.00	2782.00	2732.00			
7774.00	7466.00	3077.84	2987.00	5692.00	5546.00			
3880.47	3758.00			6110.11	5903.00			

FEATURES

Location/Qualifiers		source
1..37448	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/chromosome="6"	
	/map="6p21"	
	/clone="CGM1:DL00E8"	
	/sub_clone="UWGC:y54c194"	
	/cell_line="CGM1"	
	/clone_lib="Wash U YAC Library"	
1677..2279	/rpt_family="Alu"	repeat_region
complement(4391..4681)	/rpt_family="Alu"	repeat_region
complement(9888..10175)	/rpt_family="Alu"	repeat_region
10603..10887	/rpt_family="Alu"	repeat_region
10912..11167	/rpt_family="Alu"	repeat_region
11217..11803	/rpt_family="Alu"	repeat_region
11958..12256	/rpt_family="Alu"	repeat_region
12303..12582	/rpt_family="Alu"	repeat_region
complement(12597..12687)	/rpt_family="Alu"	repeat_region
13133..13711	/rpt_family="Alu"	repeat_region
13765..14052	/rpt_family="Alu"	repeat_region
complement(14068..14349)	/rpt_family="Alu"	repeat_region
complement(14767..15042)	/rpt_family="Alu"	repeat_region

```

repeat_region 15558..15808
/rpt_family="Alu"
repeat_region 16088..16169
/rpt_family="Alu"
repeat_region 16408..16723
/rpt_family="Alu"
repeat_region complement(16826..16898)
/rpt_family="MIR"
repeat_region complement(16962..17246)
/rpt_family="Alu"
repeat_region 18131..18416
/rpt_family="Alu"
repeat_region 18430..18718
/rpt_family="Alu"
repeat_region complement(20650..20883)
/rpt_family="Alu"
repeat_region 20973..21256
/rpt_family="Alu"
repeat_region 21612..21899
/rpt_family="Alu"
repeat_region 22077..22371
/rpt_family="Alu"
repeat_region complement(22407..22621)
/rpt_family="Alu"
repeat_region complement(22623..23251)
/rpt_family="Alu"
repeat_region complement(25158..25487)
/rpt_family="Alu"
repeat_region 27289..27574
/rpt_family="Alu"
repeat_region 27918..28216
/rpt_family="Alu"
repeat_region 28576..28865
/rpt_family="Alu"
repeat_region 29305..29563
/rpt_family="Alu"
repeat_region complement(29902..30200)
/rpt_family="Alu"
repeat_region 30514..30810
/rpt_family="Alu"
repeat_region complement(31248..31533)
/rpt_family="Alu"
repeat_region 33681..33981
/rpt_family="Alu"
repeat_region complement(34178..34470)
/rpt_family="Alu"
repeat_region 34922..35197
/rpt_family="Alu"
repeat_region 35845..36119
/rpt_family="Alu"
BASE COUNT 9686 a 9222 c 8828 g 9712 t
ORIGIN
Query Match 84.0%; Score 21; DB 9; Length 37448;
Best Local Similarity 100.0%; Pred. No. 0.004;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTGCGAGTGAGCCGAGATAA 21
|||||
Db 13639 GGTGCGAGTGAGCCGAGATAA 13659

RESULT 32
AC0000035/c
LOCUS Homo sapiens Chromosome 22q12 Cosmid Clone 1147g11, complete
DEFINITION sequence.
ACCESSION AC0000035
VERSION AC0000035.2 GI:4560484
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 38429)
Pan, H., Hartman, K., Peyrard, M., Fransson, I., Dumanski, J. P. and
Roe, B. A.
TITLE Homo sapiens Chromosome 22q12 Cosmid Clone 1147g11 In NEFH, MDR
Region
JOURNAL Unpublished
REFERENCE
AUTHORS 2 (bases 1 to 38429)
Roe, B. A.
TITLE Direct Submission
JOURNAL Submitted (21-OCT-1996) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REMARK 1147g11
REFERENCE
AUTHORS 3 (bases 1 to 38429)
Roe, B. A.
TITLE Direct Submission
JOURNAL Submitted (18-FEB-1998) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS 4 (bases 1 to 38429)
Roe, B. A.
TITLE Direct Submission
JOURNAL Submitted (11-MAR-1998) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS 5 (bases 1 to 38429)
Roe, B. A.
TITLE Direct Submission
JOURNAL Submitted (21-MAR-1998) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS 6 (bases 1 to 38429)
Roe, B. A.
TITLE Direct Submission
JOURNAL Submitted (28-MAR-1998) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS 7 (bases 1 to 38429)
Roe, B. A.
TITLE Direct Submission
JOURNAL Submitted (11-JUN-1998) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS 8 (bases 1 to 38429)
Roe, B. A.
TITLE Direct Submission
JOURNAL Submitted (01-AUG-1998) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS 9 (bases 1 to 38429)
Roe, B. A.
TITLE Direct Submission
JOURNAL Submitted (20-AUG-1998) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS 10 (bases 1 to 38429)
Roe, B. A.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-1998) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS 11 (bases 1 to 38429)
Roe, B. A.
TITLE Direct Submission
JOURNAL Submitted (30-OCT-1998) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS 12 (bases 1 to 38429)
Roe, B. A.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-1998) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

```


OK 73019, USA
 13 (bases 1 to 38429)
 Roeb.B.A.
 Direct Submission
 Submitted (13-NOV-1998) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 14 (bases 1 to 38429)
 Roeb.B.A.
 Direct Submission
 Submitted (22-NOV-1998) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 15 (bases 1 to 38429)
 Roeb.B.A.
 Direct Submission
 Submitted (04-DEC-1998) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 16 (bases 1 to 38429)
 Roeb.B.A.
 Direct Submission
 Submitted (09-APR-1999) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 17 (bases 1 to 38429)
 Pan,H., Hartman,K., Peyrard,M., Fransson,I., Dumanski,J.P. and
 Roeb.B.A.
 Direct Submission
 Submitted (01-JUN-2000) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 On Apr 9, 1999 this sequence version replaced gi:3962508.
 Because these overlapping clones came from different libraries
 there are numerous instances of insertions, deletions, and single
 nucleotide polymorphisms in the overlapping regions below
 AC000025(566c1) 27423 65851 (69554) overlaps AC000035(1147g11) 1
 38429 (0) AC000035(1147g11) 10813 38429 (0) overlaps
 AC005529(bk256d12) 1 27604 (290884) AC000035(1147g11) 11037 38429
 (0) overlaps AC005527(489d1) 1 27380 (121928).

FEATURES
 Location/Qualifiers
 I. .38429
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="22q12"
 /clone="1147g11"
 BASE COUNT 10052 a 9780 c 9922 g 8675 t
 ORIGIN

Query Match 84.0%; Score 21; DB 9; Length 38429;
 Best Local Similarity 100.0%; Pred. No. 0.0039;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCGAGTCAGCCGAGATAA 21
 |||||
 Db 25328 GGTTCGAGTCAGCCGAGATAA 25308

RESULT 33
 HSU46H11
 LOCUS
 DEFINITION Human DNA sequence from clone LLOXNC01-46H11 on chromosome X
 Contains a gene for a novel protein, complete sequence.
 ACCESSION 282254
 VERSION 282254.1 GI:1817669
 KEYWORDS RTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 38540)
 AUTHORS Dodsworth,S.
 TITLE Direct Submission

JOURNAL

Submitted (07-FEB-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquery@sanger.ac.uk
 Clon request: clonrequest@sanger.ac.uk

COMMENT

On Feb 6, 1997 this sequence version replaced gi:1781019.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em., EMBL; Sw.,
 SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP
 database can be found at

http://www.sanger.ac.uk/Projects/C.elegans/wormpep
 This sequence was generated from part of bacterial clone contigs of human
 chromosome X, constructed by the Sanger Centre Chromosome X Mapping
 Group. Further information can be found at

http://www.sanger.ac.uk/HGP/ChrX
 LLOXNC01-46H11 is from the Lawrence Livermore National Laboratory
 flow-sorted

X chromosome cosmid library LLOXNC01

VECTOR: lawrist6

This sequence is the entire insert of clone LLOXNC01-46H11.

FEATURES

Location/Qualifiers

I. .38540

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="X"

/clone="LLOXNC01-46H11"

/clone_lib="LLOXNC01"

I. .4

/note="orientation unknown"

repeat_region 27. .378

/note="L1P1A13 repeat: matches 203. .604 of consensus"

repeat_region 444. .837

/note="L1P1A13 repeat: matches -314. .84 of consensus"

repeat_region 836. .2696

/note="L1P1A12 repeat: matches -1416. .573 of consensus"

misc_feature complement(2717. .3191)

/note="match: GSS: Em:AQ807555"

repeat_region 3326. .3415

/note="L2 repeat: matches 52. .142 of consensus"

repeat_region 3428. .3544

/note="FLAM_A repeat: matches 1. .124 of consensus"

repeat_region 3556. .4338

/note="L2 repeat: matches 157. .886 of consensus"

repeat_region 4339. .4665

/note="MLF1B repeat: matches 3. .388 of consensus"

repeat_region 4666. .4997

/note="L2 repeat: matches 886. .1229 of consensus"

repeat_region 5112. .5311

/note="L2 repeat: matches 2244. .2446 of consensus"

repeat_region 5306. .6783

/note="L2 repeat: matches 1218. .2746 of consensus"

repeat_region 6798. .7353

/note="L1ME repeat: matches 5397. .5937 of consensus"

repeat_region 7354. .7667

/note="L1P1A7 repeat: matches 5827. .6141 of consensus"

repeat_region 7668. .7860

/note="L1ME repeat: matches 5186. .5398 of consensus"

repeat_region 8020. .8605

/note="L1ME3A repeat: matches 5555. .6158 of consensus"

repeat_region 9188. .9554

/note="L1R40b repeat: matches 86. .462 of consensus"

misc_feature complement(9221. .9731)

```
/note="match: GSS: Em:AQ462192"
complement(9327..9723)
/note="match: GSS: Em:AQ181119"
9734..10214
/note="match: GSS: Em:AQ468867"
9888..10332
/note="MGC repeat: matches 1..426 of consensus"
10467..10769
/note="AluX repeat: matches 1..312 of consensus"
11111..11210
/note="MIR repeat: matches 2..100 of consensus"
11485..11771
/note="AluX repeat: matches 1..288 of consensus"
11772..11807
/note="18 copies 2 mer ta 83% conserved"
12197..12286
/note="L2 repeat: matches 2207..2296 of consensus"
12496..12599
/note="L2 repeat: matches 2645..2750 of consensus"
12648..12767
/note="L2 repeat: matches 1713..1843 of consensus"
13197..13501
/note="AluSg repeat: matches 1..305 of consensus"
13592..14613
/note="L1M3 repeat: matches 4672..5721 of consensus"
14614..15487
/note="L1M4 repeat: matches 5424..6293 of consensus"
15488..15867
/note="L1M3 repeat: matches 5721..5895 of consensus"
15860..15923
/note="L1M3 repeat: matches 6044..6108 of consensus"
15926..16083
/note="L2 repeat: matches 716..875 of consensus"
16885..17646
/note="MER21B repeat: matches 1..790 of consensus"
17374..17791
/note="match: GSS: Em:AQ773679"
17787..17983
/note="L1M4 repeat: matches 7797..7977 of consensus"
17948..18574
/note="match: GSS: Em:AQ270655"
18538..18585
/note="16 copies 3 mer att 79% conserved"
18792..18927
/note="L2 repeat: matches 2528..2660 of consensus"
18928..19231
/note="AluSp repeat: matches 1..301 of consensus"
19232..19450
/note="L2 repeat: matches 2311..2528 of consensus"
complement(19468..20114)
/note="match: GSS: Em:AQ381550"
complement(19830..19905)
/note="match: STS: Em:L77851"
20128..20685
/note="match: GSS: Em:AQ802353"
complement(20418..20763)
/note="match: GSS: Em:AQ585332"
21561..22019
/note="match: GSS: Em:AQ267614"
22028..22249
/note="MER20 repeat: matches 1..218 of consensus"
22858..22957
/note="5 copies 20 mer 85% conserved"
22967..23026
/note="3 copies 20 mer 83% conserved"
22997..23152
/note="4 copies 39 mer 86% conserved"
23100..23175
/note="2 copies 38 mer 90% conserved"
24241..24404
/note="MIR repeat: matches 17..187 of consensus"
25362..25482
/note="MIR repeat: matches 21..142 of consensus"

repeat_region
25715..25937
/note="MIR repeat: matches 25..262 of consensus"
25946..26418
/note="match: GSS: Em:B44804"
27225..27406
/note="MER5A repeat: matches 11..188 of consensus"
27405..27497
/note="MIR repeat: matches 15..120 of consensus"
27566..27625
/note="30 copies 2 mer aa 80% conserved"
27852..27939
/note="MIR repeat: matches 67..150 of consensus"
28131..31922
/gene="cu46H11.1"
join(28131..28265,28627..28883,29277..29328,29516..29740)
/gene="cu46H11.1"
/product="cu46H11.1 (novel protein)"
/note="isoform 3"
/match: CDNAS: Em:AK026512"
/evidence=not_experimental
/gene="cu46H11.1"
/product="cu46H11.1 (novel protein)"
/note="isoform 4"
/match: ESTs: Em:AA334897"
/evidence=not_experimental
28423..28956
/gene="cu46H11.1"
/note="CpG island"
/evidence=not_experimental
join(28628..28766,29235..29339,29523..29684,30095..31922)
/gene="cu46H11.1"
/product="cu46H11.1 (novel protein)"
/note="isoform 1"
/match: CDNAS: Em:AK005317"
match: ESTs: Em:BG700477 Em:BG702186 Em:BG713924
Em:BG706249 Em:BG716383 Em:BG708033 Em:BG706537
Em:BG715540"
/evidence=not_experimental
29083..29168
/note="43 copies 2 mer gg 62% conserved"

repeat_region
Query Match 84.0%; Score 21; DB 9; Length 38540;
Best Local Similarity 100.0%; Pred.No.0.0039;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCGAGATAA 21
|||||
Db 11701 GGTTCAGTGAGCGAGATAA 11721

RESULT 34
AC093236 38634 bp DNA linear PRI 29-NOV-2001
LOCUS Homo sapiens chromosome 19 clone LLNLR-275D8, complete sequence.
DEFINITION AC093236
ACCESSION AC093236
VERSION AC093236.2 GI:17149385
KEYWORDS HTG
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 38634)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 38634)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (16-AUG-2001) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 38634)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
```

TITLE Direct Submission
JOURNAL Submitted (29-NOV-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Nov 29, 2001 this sequence version replaced gi:15193370. Draft sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov

Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.

FEATURES	Estimated Total Number of Errors is 0.
Location/Qualifiers	
source	1 38634

```
source
1. 38634
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
```

BASE COUNT	ORIGIN	clone="LLNLR-275D8"
9175 a	10661 c	9950 g 8848 t

Query Match 84.0%; Score 21; DB 9; Length 38634;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGGCCGAGATAA 21
|||||
Db 17592 GGTTCAGTGGCCGAGATAA 17612

RESULT 35
AC011516
LOCUS AC011516
DEFINITION Homo sapiens chromosome 19 clone LLNLF-163G11, complete sequence.
ACCESSION AC011516
40014 bp DNA linear PRI 26-JUN-2001

VERSION AC011516.3 GT:14550303
KEYWORDS HTG.
SOURCE Homo sapiens.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 40014)

REFERENCE

DOE Joint Genome Institute and Stanford Human Genome Center.
 (pages 1 to 40014)
 AUTHORS
 TITLE
 Direct Submission

JOURNAL
REFERENCE
AUTHORS
TITLE
UNPUBLISHED
2 (bases 1 to 40014)
DOE Joint Genome Institute.

JOURNAL Direct Submission
REFERENCE Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 40014)

AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 TOTAL 1

Submitted (26-JUN-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

On Jun 26, 2001 this sequence version replaced gi:7690152.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov

www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu

FEATURES

ERRORS	Location/Qualifiers
source	1. .40014

```

ASE COUNT      12075      869
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="LLNLF-163G11"

```

Query Match 84.0%; Score 21; DB 9; Length 40014;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 1 GGTTGCAGTCAGCCGAGATAA 21
|||||
Db 17874 GGTTGCAGTCAGCCGAGATAA 17894

RESULT 36					
HSB42B1					
LOCUS	HSB42B1	40943 bp	DNA	linear	PRI 12-DEC-1999
DEFINITION	Human DNA sequence from clone SC2CB-42B1 on chromosome 22 Contains STSs and GSSs, complete sequence.				

ACCESSION	Z75890	
VERSION	Z75890.1	GI:1430781
KEYWORDS	ntc	

SOURCE	Homo sapiens.
ORGANISM	Homo sapiens.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 40943)
 Hunt S

TITLE	Direct Submission
JOURNAL	Submitted (08-DEC-1999) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: clonerequests@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
COMMENT	

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/Worm>.

http://www.sanger.ac.uk/projects/c elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 22, constructed by the Sanger Centre Chromosome 22 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr22> sc22054221

SC22CB 42b1 is from the human chromosome 22-specific cosmid library (SC22CB) constructed at the Sanger Centre by Mark Ross and Cordelia Langford.

VECTOR: lawrist16
This sequence is the entire insert of clone sc3008-12m1

FEATURES	Location/Qualifiers
source	1. .40943

```
/organism="Homo sapiens"  
/db_xref="taxon:9606"
```

```

/chromosome="22"
/clone="SC22CB-42B1"
/clone_lib="naaa"

```

```
repeat_region
1269.1565
/crime_id="SC22QB"
/notes="Alusq report"
```

```
repeat_region
/note="Alu repeat: matches 1 100 of consensus"
1269..1562
/note="Alu repeat: matches 1 100 of consensus"
```

```
repeat_region
2199..2359
      repeat: matches 1..100% of consensus"
/note="AluX repeat: matches 135..206 of consensus"
```

```

repeat_region
complement(2204, .2366)
/note="Alu repeat: matches 138, .308 of consensus"

```

```
repeat_region 2360..2655
                /note="AluY repeat: matches 1..298 of consensus"
repeat_region 2360..2655
```

```
repeat_region
complement(2367..2655)
/note="Alu repeat: matches 1..308 of consensus"
2556 2782
repeat_region
```

```
repeat_region
2000..12782
/note="AluX repeat: matches 9..135 of consensus"
complement(2657..2703)
repeat_region
```

```
repeat_region
2823..3122
/note="Alu repeat: matches 1..150 of consensus"
complement(2057..2192)
```

```
repeat_region /note="AluX repeat: matches 1. .299 of consensus"  
complement(2830. .3122)  
repeat_region /note="Alu repeat: matches 1. .308 of consensus"  
3510. .3639  
repeat_region /note="MIR repeat: matches 20. .144 of consensus"  
3987. .4282  
misc_feature /note="L1 repeat: matches 4735. .5039 of consensus"  
5043. .5528  
repeat_region /note="match: GSS: Em:AQ560214"  
5469. .5565  
repeat_region /note="L1MEC repeat: matches 2340. .2085 of consensus"  
5584. .5653  
repeat_region /note="35 copies 2 mer ct 70 conserved"  
5874. .6994  
misc_feature /note="L1MEC repeat: matches 1127. .2299 of consensus"  
6653. .7277  
repeat_region /note="match: GSS: Em:AQ371546"  
7258. .7629  
repeat_region /note="L1MEC repeat: matches 271. .669 of consensus"  
7665. .7776  
repeat_region /note="MER3 repeat: matches 1. .113 of consensus"  
7831. .8164  
repeat_region /note="L1MCb repeat: matches 137. .458 of consensus"  
8263. .8376  
repeat_region /note="HERVL repeat: matches 791. .900 of consensus"  
8377. .8497  
repeat_region /note="MER20 repeat: matches 98. .218 of consensus"  
8506. .8858  
repeat_region /note="L1MC4 repeat: matches 6243. .6607 of consensus"  
8909. .8994  
repeat_region /note="43 copies 2 mer at 74 conserved"  
8910. .8993  
repeat_region /note="14 copies 6 mer tatata 75 conserved"  
8997. .10192  
repeat_region /note="L1MC4 repeat: matches 6605. .7977 of consensus"  
11506. .11825  
repeat_region /note="AluSg repeat: matches 1. .310 of consensus"  
11506. .11805  
repeat_region /note="Alu repeat: matches 1. .308 of consensus"  
11835. .11902  
repeat_region /note="34 copies 2 mer tt 72 conserved"  
12297. .12712  
misc_feature /note="L2 repeat: matches 2370. .2749 of consensus"  
complement(12728. .13109)  
repeat_region /note="match: GSS: Em:AQ421066"  
13814. .13959  
repeat_region /note="MIR repeat: matches 50. .203 of consensus"  
14346. .14472  
repeat_region /note="L2 repeat: matches 2342. .2482 of consensus"  
14691. .14812  
repeat_region /note="MIR repeat: matches 127. .252 of consensus"  
14763. .14820  
repeat_region /note="L2 repeat: matches 2646. .2702 of consensus"  
15030. .15228  
repeat_region /note="MIR repeat: matches 6. .210 of consensus"  
15849. .16240  
repeat_region /note="L2 repeat: matches 2325. .2700 of consensus"  
16325. .16374  
repeat_region /note="L2 repeat: matches 2144. .2194 of consensus"  
17563. .17790  
repeat_region /note="MIR repeat: matches 1. .251 of consensus"  
18915. .19244  
misc_feature /note="MER2 repeat: matches 13. .344 of consensus"  
19333. .19777  
repeat_region /note="match: GSS: Em:AQ353457"  
19734. .19917  
repeat_region /note="MER58A repeat: matches 28. .222 of consensus"  
20077. .20112  
repeat_region /note="MIR repeat: matches 84. .119 of consensus"  
21875. .22039  
repeat_region /note="L2 repeat: matches 2579. .2745 of consensus"  
22033. .22261  
/note="L2 repeat: matches 1838. .2050 of consensus"
```

```
repeat_region 22262. .22589  
/note="AluX repeat: matches 1. .309 of consensus"  
22262. .22552  
repeat_region /note="Alu repeat: matches 1. .308 of consensus"  
22570. .23052  
repeat_region /note="L2 repeat: matches 1288. .1838 of consensus"  
23840. .24130  
repeat_region /note="Alu repeat: matches 1. .308 of consensus"  
23842. .24136  
repeat_region /note="AluSg repeat: matches 3. .299 of consensus"  
24154. .24365  
misc_feature /note="MIR11 repeat: matches 198. .408 of consensus"  
24337. .24562  
repeat_region /note="match: GSS: Em:T91723"  
24434. .24553  
repeat_region /note="L2 repeat: matches 1157. .1276 of consensus"  
24580. .24871  
repeat_region /note="Alu repeat: matches 1. .308 of consensus"  
24584. .24879  
repeat_region /note="AluSg repeat: matches 5. .301 of consensus"  
24950. .25190  
repeat_region /note="L1 repeat: matches 4576. .4812 of consensus"  
25198. .25489  
repeat_region /note="Alu repeat: matches 1. .308 of consensus"  
25200. .25502  
repeat_region /note="AluJo repeat: matches 3. .305 of consensus"  
25509. .26154  
repeat_region /note="L1MA4 repeat: matches 5461. .6139 of consensus"  
26178. .26217  
repeat_region /note="20 copies 2 mer at 97 conserved"  
26180. .26215  
misc_feature /note="6 copies 6 mer atatat 100 conserved"  
complement(26608. .27122)  
repeat_region /note="match: GSS: Em:AQ594970"  
27081. .27325  
repeat_region /note="L1MC/D repeat: matches 5330. .5587 of consensus"  
27339. .27368  
repeat_region /note="15 copies 2 mer gt 100 conserved"  
27605. .27745  
repeat_region /note="L2 repeat: matches 2560. .2709 of consensus"  
29581. .29746  
repeat_region /note="L2 repeat: matches 2587. .2750 of consensus"  
29839. .30025  
repeat_region /note="L2 repeat: matches 2121. .2301 of consensus"  
complement(30157. .30263)  
repeat_region /note="Alu repeat: matches 28. .153 of consensus"  
30162. .30262  
repeat_region /note="FLAM_A repeat: matches 28. .128 of consensus"  
30497. .30538  
repeat_region /note="7 copies 6 mer ctctct 85 conserved"  
30506. .30537  
repeat_region /note="16 copies 2 mer tc 90 conserved"  
31615. .31765  
repeat_region /note="MIR repeat: matches 98. .262 of consensus"  
32163. .32488  
repeat_region /note="L2 repeat: matches 2059. .2491 of consensus"  
32499. .32862  
misc_feature /note="THE1B repeat: matches 1. .364 of consensus"  
complement(32876. .33457)  
/note="match: GSS: Em:AQ582892"  
complement(32969. .33457)
```

Query Match 84.0%; Score 21; DB 9; Length 40943;
Best Local Similarity 100.0%; Pred. NO. 0.0039;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTGCAGTGAGCCGAGATAA 21
|||||
Db 22478 GGTTGCAGTGAGCCGAGATAA 22498

RESULT 37
AL355881/c

LOCUS AL355881 42304 bp DNA linear PRI 01-MAR-2001
 DEFINITION Human DNA sequence from clone RP11-332B14 on chromosome 6, complete
 sequence.
 ACCESSION AL355881
 VERSION AL355881.15 GI:13234900
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Hall, R.
 Direct Submission
 Submitted (01-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On Mar 5, 2001 this sequence version replaced gi:12750841.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: EM: EMBL; SW:
 SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP
 database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
 was generated from part of bacterial clone contigs This sequence
 Chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
 Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr6>
 RP11-332B14 is from the library RPC1-11.2 constructed by the group
 of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pBACe3.6
 IMPORTANT: This sequence is not the entire insert of clone
 RP11-332B14. It may be shorter because we sequence overlapping
 sections only once, except for a 100 base overlap.
 The true left end of clone RP3-528L19 is at 42205 in this sequence.
 The true right end of clone RP1-188K17 is at 100 in this sequence.
 Location/Qualifiers
 1..42304
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="6"
 /clone="RP11-332B14"
 /clone_lib="RPC1-11.2"
 612..977
 /note="L1MC1 repeat: matches 5902..6332 of consensus"
 1652..1724
 /note="FLAM_A repeat: matches 52..129 of consensus"
 1725..1908
 /note="ALUSg/x repeat: matches 81..280 of consensus"
 1976..2003
 /note="BC200 repeat: matches 1..26 of consensus"
 2004..2096
 /note="FLAM_C repeat: matches 33..125 of consensus"
 2167..2306
 /note="MIR repeat: matches 110..259 of consensus"
 2465..2741
 /note="ALUJo repeat: matches 1..284 of consensus"
 2985..3116
 /note="ALUSg/x repeat: matches 1..125 of consensus"
 3154..3309
 /note="MIR repeat: matches 26..187 of consensus"

repeat_region 3320..3630
 /note="AluY repeat: matches 1..310 of consensus"
 3856..4159
 /note="AluJo repeat: matches 1..299 of consensus"
 4390..4703
 /note="AluSg repeat: matches 1..313 of consensus"
 4766..4817
 /note="26 copies 2 mer ta 88% conserved"
 4821..4856
 /note="18 copies 2 mer ta 94% conserved"
 5041..5154
 /note="AluDb repeat: matches 1..108 of consensus"
 5186..5496
 /note="AluSx repeat: matches 1..310 of consensus"
 5539..5840
 /note="AluSx repeat: matches 1..295 of consensus"
 6385..6462
 /note="39 copies 2 mer tt 78% conserved"
 6517..6804
 /note="AluJo repeat: matches 1..290 of consensus"
 7656..7954
 /note="AluY repeat: matches 1..301 of consensus"
 8630..8896
 /note="AluSx repeat: matches 42..309 of consensus"
 9126..9439
 /note="AluSg repeat: matches 1..308 of consensus"
 10248..10538
 /note="AluSx repeat: matches 2..297 of consensus"
 10832..11041
 /note="L1PA14 repeat: matches 5941..6148 of consensus"
 11049..11353
 /note="AluSx repeat: matches 1..305 of consensus"
 11354..11665
 /note="AluSx repeat: matches 1..299 of consensus"
 12034..12123
 /note="L2 repeat: matches 1326..1414 of consensus"
 12264..12412
 /note="MIR repeat: matches 105..261 of consensus"
 12470..12554
 /note="L2 repeat: matches 2448..2518 of consensus"
 12555..12834
 /note="AluYs5 repeat: matches 35..311 of consensus"
 12835..12921
 /note="L2 repeat: matches 2357..2448 of consensus"
 13051..13169
 /note="FLAM_C repeat: matches 2..120 of consensus"
 13213..13515
 /note="AluSx repeat: matches 1..303 of consensus"
 13714..13950
 /note="MER45A repeat: matches 1..236 of consensus"
 14007..14320
 /note="AluSx repeat: matches 1..311 of consensus"
 14438..14558
 /note="AluDb repeat: matches 1..121 of consensus"
 15342..15413
 /note="L2 repeat: matches 2632..2703 of consensus"
 15982..16284
 /note="AluDb repeat: matches 1..311 of consensus"
 16309..16544
 /note="Alu repeat: matches 6..293 of consensus"
 18245..18349
 /note="AluYa8 repeat: matches 100..210 of consensus"
 18362..18670
 /note="AluSx repeat: matches 1..310 of consensus"
 19118..19431
 /note="AluJo repeat: matches 1..303 of consensus"
 19656..19789
 /note="FLAM_C repeat: matches 1..132 of consensus"
 19997..20298
 /note="AluSx repeat: matches 1..302 of consensus"
 20725..21017
 /note="AluJo repeat: matches 1..305 of consensus"
 21387..21452

repeat_region /note="Alu repeat: matches 237. .302 of consensus"
21551. .21586
repeat_region /note="Wln2E repeat: matches 110. .143 of consensus"
21587. .21885
repeat_region /note="Alusx repeat: matches 5. .303 of consensus"
21886. .22127
repeat_region /note="MLT2E repeat: matches 143. .397 of consensus"
22305. .22403
repeat_region /note="HY3 repeat: matches 2. .100 of consensus"
22421. .22716
repeat_region /note="Alusx repeat: matches 1. .296 of consensus"
22778. .23115
repeat_region /note="AluJb repeat: matches 1. .312 of consensus"
23297. .23606
repeat_region /note="Alusx repeat: matches 1. .308 of consensus"
23607. .23687
repeat_region /note="LIPB2 repeat: matches 5890. .5973 of consensus"
23688. .23995
repeat_region /note="Alusx repeat: matches 1. .310 of consensus"
23996. .24388
repeat_region /note="LIPB2 repeat: matches 5541. .5890 of consensus"
24389. .24708
repeat_region /note="Alusx repeat: matches 1. .301 of consensus"
24709. .24958
repeat_region /note="LIPB2 repeat: matches 5281. .5541 of consensus"
25450. .25683
repeat_region /note="AluJb repeat: matches 1. .291 of consensus"
25691. .25978
repeat_region /note="Aluy repeat: matches 1. .302 of consensus"
26205. .26483
repeat_region /note="Alusx repeat: matches 1. .301 of consensus"
26556. .26671
repeat_region /note="58 copies 2 mer tt 59% conserved"
28625. .28939
repeat_region /note="MER33 repeat: matches 8. .323 of consensus"
28978. .29078
repeat_region /note="AluJo repeat: matches 43. .146 of consensus"
29079. .29386
repeat_region /note="Alusq repeat: matches 1. .307 of consensus"
29387. .29565
repeat_region /note="AluJo repeat: matches 146. .312 of consensus"
29575. .29635
repeat_region /note="MERS1B repeat: matches 466. .527 of consensus"
29875. .30078
repeat_region /note="L2 repeat: matches 2192. .2408 of consensus"
30074. .30168
repeat_region /note="L2 repeat: matches 2654. .2750 of consensus"
30491. .30609
repeat_region /note="MERS7-internal repeat: matches 7148. .7267 of
consensus"
30991. .31288
repeat_region /note="Alusx repeat: matches 1. .296 of consensus"
31415. .31453
repeat_region /note="MIR repeat: matches 106. .144 of consensus"
84.0%; Score 21; DB 9; Length 42304;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGATAA 21
|||||
Db 33829 GGTTCAGTGCAGCCGAGATAA 33809

RESULT 38
HSU63963
DEFINITION Human CSF-1 receptor (FMS) gene, complete cds, and (SMF) gene,
partial cds.
ACCESSION U63963 X14720
VERSION U63963.1 GI:1915975
KEYWORDS
SOURCE Homo sapiens.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Hampe,A., Shamoon,B.M., Gobet,M., Sherr,C.J. and Galibert,F.
TITLE Nucleotide sequence and structural organization of the human FMS
proto-oncogene
Oncogene Res. 4 (1), 9-17 (1989)
JOURNAL 892339490
MEDLINE 2524025
PUBMED
REFERENCE 2 (bases 1 to 42874)
AUTHORS Gastier,J.M., Brody,T., Pulido,J.C., Businga,T., Sundén,S., Hu,X.,
Maitra,S., Buetow,K.H., Murray,J.C., Sheffield,V.C., Boguski,M.,
Duyk,G.M. and Hudson,T.J.
TITLE Development of a screening set for new (CAG/CTG)_n dynamic mutations
Genomics 32 (1), 75-85 (1996)
JOURNAL 96230328
MEDLINE 8786123
PUBMED
REFERENCE 3 (bases 1 to 42874)
AUTHORS Andre,C., Hampe,A., Lachaux,P., Martin,E., Wang,X.P., Manus,V.,
Hu,W.X. and Galibert,F.
TITLE Sequence analysis of two genomic regions containing the KIT and the
FMS receptor tyrosine kinase genes
Genomics 39 (2), 216-226 (1997)
JOURNAL 97179223
MEDLINE 9027509
PUBMED
REFERENCE 4 (bases 1 to 42874)
AUTHORS Hampe,A.
TITLE Direct Submission
Submitted (17-JUL-1996) UPR41, CNRS, Avenue Leon Bernard, Rennes
35000, France
COMMENT On Apr 12, 1997 this sequence version replaced gi:30238.
FEATURES
source 1..42874
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5q33.3-q34"
/tissue_type="placenta"
gene 1..35016
/gene="FMS"
intron <1..1735
/gene="FMS"
/number=1
repeat_region 308..591
/rpt_family="Alu J"
/rpt_type-dispersed
repeat_region complement(719..902)
/rpt_family="MER32"
repeat_region 1264..1347
/rpt_family="Alu J"
/rpt_type-dispersed
mRNA join(1736..1964,7314..7571,8002..8286,10090..10226,
10903..11062,14843..15035,17766..17881,18035..18155,
18274..18464,20007..20122,26478..26604,26732..26836,
27354..27464,28464..28626,30734..30822,30942..31039,
31985..32107,32189..32300,32991..33090,33897..34005,
34103..>34258)
/gene="FMS"
exon /product="CSF-1 receptor"
1736..1964
/gene="FMS"
/number=2
CDS join(1916..1964,7314..7571,8002..8286,10090..10226,
10903..11062,14843..15035,17766..17881,18035..18155,
18274..18464,20007..20122,26478..26604,26732..26836,
27354..27464,28464..28626,30734..30822,30942..31039,
31985..32107,32189..32300,32991..33090,33897..34005,
34103..34258)
/gene="FMS"
/function="growth factor receptor"
/function="tyrosine kinase receptor"
/note="c-fms proto-oncogene; transmembrane glycoprotein"

```
/codon_start=1
/product="CSF-1 receptor"
/protein_id="AAB51696.1"
/db_xref="GI:1915976"
/translation="MGPGVLLLLVATAMHGQIPVIEPSPVPELVKPGAIIVTLRCVG
NGSVEWDGPPSPHWTLYSDGSSSILSNNTAFTGTTCRTEPDPLGSSAAIHLYVK
DPAPRNVLAEQVYVFEODALLPCLLTPVLEAGVSLVRGRPLMRHTNYSFSPWH
GFTIIRAKFIQSODYCCSALMGGRKVMISIRLKVKVIFGPPALTLVPAELVIRGE
AAQYASASVDNFDVFOHNTKLAIPQSDPHNNYOKVLTNLQDVDFPAGNY
SCVASNVQGHSTSMFFRVESAYLNLSEQNLQEVTVGEGNLKLVWYVAPGLQGF
NWTVLGPFSDHOPEPKLANATTKTYVHTFTLSLIPKPSBAGKISFLARNPGGWRAL
TFELTLRYPEVSVLWTFINGSTLLCAAGSPQPNVTWLOCSGHTDRDAQLQVW
DDPYEVLQEPFPHKVTQSLITVEILEHNTYICRAHNSVSGSWAFTIPISAGATH
PPDFLETPYVACMSIMALLLLLLLYLKQPKYQVRWKIIESTEGSYTFIDP
TQLYNEKWEFFRNNOFGKTLGAGAFKVVVATAFGLGEDAVLKAVKMLKSTAH
DEKALMSCLKIMSHLGHENIVNLLGACTHGGPVLVITRYCCYGLDLNFTLRARBM
LGPSSLPQDPEGVDYKNIHLKKYVRDGSFGQVDIYVEMRPVSTSNDSFSQ
DLKEDGRPLELRLDLHFSOVAQMAFLASKNCLHRDVAARNVLLTNGHVAKIGDFG
LARDIMDNGNYIVKGNARLPVKWAPESIFDCVITYVSDVMSYGLLWEIFSLGLNPY
PGIIVNSKFKLVKDGQMAQPAFAPKNYISIMQACWALEPHTPTFOQICSFQEOA
QEURRERYTNLPSSRSRGGSGSSSELESSEHLLTCCQGDIAQLPLQPNNYQFC
"
1965..7313
/gene="FMS"
/number=2
complement(3069..3346)
/rpt_family="Alu Sx"
/rpt_type-dispersed
3592..3872
/rpt_family="Alu Sx"
/rpt_type-dispersed
3872..3995
/note="(CCTT)19(CTTT)12 repeat"
/rpt_type-tandem
/rpt_unit=CTTT
complement(4009..4284)
/rpt_family="Alu Sx"
/rpt_type-dispersed
4938..4989
/note="(GT)26 repeat"
/rpt_type-tandem
/rpt_unit=GT
7314..7571
/gene="FMS"
/number=3
7572..8001
/gene="FMS"
/number=3
8002..8286
/gene="FMS"
/number=4
8287..10089
/gene="FMS"
/number=4
10090..10226
/gene="FMS"
/number=5
10227..10902
/gene="FMS"
/number=5
10903..11062
/gene="FMS"
/number=6
11063..14842
/gene="FMS"
/number=6
12099..12173
/rpt_family="Alu Sx"
/rpt_type-dispersed
13861..14008
/rpt_family="Alu Sc"
/rpt_type-dispersed
```

```
exon 14843..15035
/gene="FMS"
/number=7
15036..17765
/gene="FMS"
/number=7
repeat_region
complement(15848..16124)
/rpt_family="Alu Sx"
/rpt_type-dispersed
17766..17881
/gene="FMS"
/number=8
17882..18034
/gene="FMS"
/number=8
18035..18155
/gene="FMS"
/number=9
18156..18273
/gene="FMS"
/number=9
18274..18464
/gene="FMS"
/number=10
18465..20006
/gene="FMS"
/number=10
repeat_region
19098..19143
/note="repeat sequence I"
19915..19967
/note="repeat sequence I"
20007..20122
/gene="FMS"
/number=11
20123..26477
/gene="FMS"
/number=11
20366..20480
/rpt_family="Alu J"
/rpt_type-dispersed
complement(22534..22568)
/rpt_family="IVR"
22759..22799
/note="repeat sequence II"
complement(24034..24217)
/rpt_family="MER13"
24460..24499
repeat_region
```

Query Match 84.0%; Score 21; DB 9; Length 42874;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGATAA 21
|||||
DB 37789 GGTTCAGTGCAGCCGAGATAA 37809

RESULT 39
AL357952/c

LOCUS

DEFINITION Human DNA sequence from clone RP11-479K1 on chromosome 1 Contains

STSS and GSSs, complete sequence.

ACCESSION AL357952

VERSION AL357952.7 GI:9863832

KEYWORDS HTG.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Wallis,J.

TITLE Direct Submission

JOURNAL Submitted (03-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk

On Aug 21, 2000 this sequence version replaced gi:9801404. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/Projects/C/elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chrl>

IMPORTANT: This sequence is not the entire insert of clone RP11-479K1 it may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true right end of clone RP11-479K1 is at 43456 in this sequence. The true right end of clone RP11-122C9 is at 100 in this sequence. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated repeat sequence elements. Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. RP11-479K1 is from the library RPC1-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACE3.6.

```

/misc_feature      /note="10 copies 4 mer tctc 90% conserved"
/misc_feature      complement(26376..26897)
/misc_feature      /note="match: GSS: Em:A0676452"
/misc_feature      complement(28822..29342)
/misc_feature      /note="match: GSS: Em:A0741080"
/misc_feature      30351..30791
/misc_feature      /note="match: GSS: Em:A0442484"
/repeat_region    30772..31126
/repeat_region    /note="THE1A repeat: matches 1..354 of consensus"
/repeat_region    31127..32702
/repeat_region    /note="THE1A internal repeat: matches 1..1580 of consensus"
/repeat_region    32703..33069
/repeat_region    /note="THE1A repeat: matches 1..354 of consensus"
/misc_feature      34363..34366
/misc_feature      /note="Single clone region region. Assembly confirmed by restriction digest data."
/repeat_region    35077..35597
/repeat_region    /note="L1MAA repeat: matches 5785..6300 of consensus"
/repeat_region    35598..35653
/repeat_region    /note="28 copies 2 mer ta 80% conserved"
/repeat_region    35663..35690
/repeat_region    /note="14 copies 2 mer ca 96% conserved"
/repeat_region    35666..35689
/repeat_region    /note="6 copies 4 mer acac 100% conserved"
/repeat_region    35692..35787
/repeat_region    /note="L1MAAA repeat: matches 5688..5782 of consensus"
/repeat_region    35789..35974
/repeat_region    /note="L1 repeat: matches 3191..3373 of consensus"
/repeat_region    35977..42134
/repeat_region    /note="L1PAA repeat: matches 1..6146 of consensus"
/repeat_region    42136..43189
/repeat_region    /note="L1 repeat: matches 2124..3209 of consensus"
BASE COUNT      12269 a 8310 c 8428 g 14449 t
ORIGIN

```

Query Match	84.0%	Score 21;	DB 9;	Length 43456;
Best Local Similarity	100.0%;	Pred. No. 0.0039;		
Matches 21;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GTTGTCAGTGAGCCGAGATAA 21		
Db	4483	GTTGTCAGTGAGCCGAGATAA 4463		
RESULT 40				
AC068187		44986 bp	DNA	linear
LOCUS				
DEFINITION	AC068187	Homo sapiens chromosome 4 clone RP11-360F5 map 4, LOW-PASS SEQUENCE SAMPLING.		
ACCESSION	AC068187			
VERSION	1	GI:7670230		
KEYWORDS	HTG; HTGS, PHASE0.			
SOURCE		Homo sapiens.		
ORGANISM		Homo sapiens		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE		1 (bases 1 to 44986)		
AUTHORS	Birren,B., Linton,L., Nusbaum,C. and Lander,E.			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 44986)			
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F., Boguslavsky,L., Bouck-Galter, B., Brown,A., Burkett,G., Campiano,A., Castile,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Glinde,S., Goyette,M., Graham,L., Grand-pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marcuis,N.,			

McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
Meldrum, J., Meneis, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rotman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE
JOURNAL

COMMENT

Submitted (29-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L9598
Center clone name: 360_F_5

* NOTE: This record contains 51 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
801 900: contig of 800 bp in length
901 1676: contig of 100 bp
1677 1776: gap of 100 bp
1777 2556: contig of 786 bp in length
2557 2656: gap of 100 bp
2657 3449: contig of 793 bp in length
3450 3549: gap of 100 bp
3550 4345: contig of 796 bp in length
4346 4445: gap of 100 bp
4446 5207: contig of 762 bp in length
5208 5307: gap of 100 bp
5308 6082: contig of 775 bp in length
6083 6182: gap of 100 bp
6183 6980: contig of 798 bp in length
6981 7080: gap of 100 bp
7081 7873: contig of 793 bp in length
7874 7973: gap of 100 bp
7974 8762: contig of 789 bp in length
8763 8862: gap of 100 bp
8863 9641: contig of 779 bp in length
9642 9741: gap of 100 bp
9742 10526: contig of 785 bp in length
10527 10626: gap of 100 bp
10627 11391: contig of 765 bp in length
11392 11491: gap of 100 bp
11492 12275: contig of 784 bp in length
12276 13375: gap of 100 bp
13376 13169: contig of 794 bp in length
13170 13269: gap of 100 bp
13270 14043: contig of 774 bp in length
14044 14143: gap of 100 bp
14144 14932: contig of 789 bp in length
14933 15032: gap of 100 bp
15033 15817: contig of 785 bp in length
15818 15917: gap of 100 bp
15918 16690: contig of 773 bp in length

FEATURES
source

1. .44986
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-360F5"
/clone.lib="RPI-11 Human Male BAC"

BASE COUNT

9880 a 9802 c 8943 g 11118 t 5243 others

16591 16790: gap of 100 bp
16791 17583: contig of 793 bp in length
17584 17683: gap of 100 bp
17684 18469: contig of 786 bp in length
18470 18569: gap of 100 bp
18570 19361: contig of 792 bp in length
19362 19461: gap of 100 bp
19462 20254: contig of 793 bp in length
20255 20354: gap of 100 bp
20355 21145: contig of 791 bp in length
21146 21245: gap of 100 bp
21246 22048: contig of 803 bp in length
22049 22148: gap of 100 bp
22149 22940: contig of 792 bp in length
22941 23040: gap of 100 bp
23041 23828: contig of 788 bp in length
23829 23928: gap of 100 bp
23929 24722: contig of 794 bp in length
24723 24822: gap of 100 bp
24823 25608: contig of 786 bp in length
25609 25708: gap of 100 bp
25709 26475: contig of 767 bp in length
26476 26575: gap of 100 bp
26576 27372: contig of 797 bp in length
27373 27472: gap of 100 bp
27473 28235: contig of 763 bp in length
28236 28335: gap of 100 bp
28336 29118: contig of 783 bp in length
29119 29218: gap of 100 bp
29219 30000: contig of 782 bp in length
30001 30100: gap of 100 bp
30101 30879: contig of 779 bp in length
30880 30979: gap of 100 bp
30980 31780: contig of 801 bp in length
31781 31880: gap of 100 bp
31881 32652: contig of 772 bp in length
32653 32752: gap of 100 bp
32753 33539: contig of 787 bp in length
33540 33639: gap of 100 bp
33640 34424: contig of 785 bp in length
34425 34524: gap of 100 bp
34525 35282: contig of 758 bp in length
35283 35382: gap of 100 bp
35383 36137: contig of 755 bp in length
36138 36237: gap of 100 bp
36238 37026: contig of 789 bp in length
37027 37126: gap of 100 bp
37127 37898: contig of 772 bp in length
37899 37998: gap of 100 bp
37999 38783: contig of 785 bp in length
38784 38883: gap of 100 bp
38884 39673: contig of 790 bp in length
39674 39773: gap of 100 bp
39774 40535: contig of 762 bp in length
40536 40635: gap of 100 bp
40636 41424: contig of 789 bp in length
41425 41524: gap of 100 bp
41525 42294: contig of 770 bp in length
42295 42394: gap of 100 bp
42395 43204: contig of 810 bp in length
43205 43304: gap of 100 bp
43305 44105: contig of 801 bp in length
44106 44205: gap of 100 bp
44206 44986: contig of 781 bp in length.

ORIGIN

Query Match 84.0%; Score 21; DB 2; Length 44986;
 Best Local Similarity 100.0%; Pred. No. 0.0039;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGGCCGAGATAA 21
 |||||
 DB 35085 GGTTCAGTGGCCGAGATAA 35105

RESULT 41
 AC005387 46213 bp DNA linear PRI 05-AUG-1998
 DEFINITION Homo sapiens chromosome 19, cosmid R29473, complete sequence.
 ACCESSION AC005387
 VERSION AC005387.1 GI:3395417
 KEYWORDS HTG.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Lamerdin,J.E., McCready,P.M., Skowronski,E., Adamson,A.W.,
 Burkhart-Schultz,K., Gordon,L., Kyle,A., Ramirez,M., Stilwagen,S.,
 Phan,H., Velasco,N., Do,L., Redala,W., Terry,A., Garnes,J.,
 Dangnan,L., Poundstone,P., Christensen,M., Georgescu,A., Avila,J.,
 Liu,S., Attix,C., Andreise,R., Frankheim,M., Amico-Keller,G.,
 Coefield,J., Duarte,S., Lucas,S., Bruce,R., Thomas,P., Quan,G.,
 Krommiller,B., Arellano,A., Montgomeri,M., Ow,D., Nolan,M.,
 Truong,S., Kobayashi,A., Olsen,A.S. and Carrano,A.V.
 Sequence analysis of an ~1 Mb region containing the MEF2B gene in
 19p12

TITLE Unpublished
 JOURNAL 2 (bases 1 to 46213)
 REFERENCE Lamerdin,J.E.
 AUTHORS Direct Submission
 JOURNAL Submitted (05-AUG-1998) Joint Genome Institute, Lawrence Livermore
 National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
 COMMENT Map and sequence oriented from p telomere to centromere. Cosmid
 R29473 is the telomeric end of the current sequencing tiling path.
 R29473 overlaps cosmid R26645 (AC005253) to the right from bases
 37,603 to 46,213. Additional chr 19 map and sequence information
 are available at: <http://www-bio.llnl.gov/genome/genome.html>.

FEATURES
 source
 1..46213
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="19"
 /map="19p12 between UBA52 and D19S451"
 /clone="R29473"
 /cell_line="5HL2-B"
 /clone_lib="LL19NC03 R chromosome 19 cosmid library"
 /note="LL19NC03 cosmid library constructed at LLNL from
 flow-sorted chromosomes from hybrid 5HL2-B, which carries
 chromosome 19 as its only human chromosome."
 19..179
 /rpt_family="FRAM"
 repeat_region 514..720
 repeat_region 514..720
 /rpt_family="MER58A"
 repeat_region 1405..1522
 /rpt_family="(CA)n"
 repeat_region 1951..2252
 /rpt_family="AluSx"
 repeat_region 2534..2703
 /rpt_family="LIME"
 repeat_region 2704..2849
 /rpt_family="AluSx"
 repeat_region 2850..3144
 /rpt_family="AluSx"
 repeat_region 3145..3274
 /rpt_family="AluSx"
 repeat_region 3276..3459
 /rpt_family="AluSx"

repeat_region 3601..3735
 /rpt_family="LIME"
 repeat_region 3739..4038
 /rpt_family="AluJo"
 repeat_region 4039..4209
 /rpt_family="AluSg"
 repeat_region 4220..4451
 /rpt_family="AluJo"
 repeat_region 5514..5814
 /rpt_family="AluJb"
 repeat_region 6096..6397
 /rpt_family="AluY"
 repeat_region 8358..8479
 /rpt_family="AluY"
 repeat_region complement(9124..9414)
 /rpt_family="LINE2"
 repeat_region 9454..9770
 /rpt_family="AluSx"
 repeat_region 9804..10070
 /rpt_family="AluJo"
 repeat_region 10083..10370
 /rpt_family="AluSx"
 repeat_region 10374..10676
 /rpt_family="AluJo"
 repeat_region 10917..11115
 /rpt_family="AluSx"
 repeat_region 11391..11692
 /rpt_family="AluY"
 repeat_region complement(12361..12422)
 /rpt_family="AluSx"
 repeat_region complement(13493..13545)
 /rpt_family="MIR"
 repeat_region complement(13787..13921)
 /rpt_family="GC-rich"
 mRNA complement(<13787..13933)
 gene /product="Human ELL mRNA"
 complement(<13787..13921)
 /gene="ELL"
 CDS /note="eleven-nineteen lysine-rich leukemia gene"
 complement(<13787..13921)
 /gene="ELL"
 /note="ELL PROTEIN"
 /codon_start=1
 /product="ELL_HUMAN"
 /protein_id="BAC28752.1"
 /db_xref="GI:3395418"
 /translation="MAALKEDRSYGLSCGRVDSGVSVFHVKLTDLSALRAFTSYRAR
 Q"
 complement(13937..13999)
 /rpt_family="GC-rich"
 repeat_region complement(14689..14854)
 /rpt_family="AluJb"
 repeat_region complement(14872..15170)
 /rpt_family="AluSx"
 repeat_region complement(15173..15479)
 /rpt_family="AluSg"
 repeat_region complement(15531..15572)
 /rpt_family="LIME6"
 repeat_region complement(15580..15877)
 /rpt_family="AluSx"
 repeat_region 16012..16324
 /rpt_family="AluJo"
 repeat_region 16345..16626
 /rpt_family="AluSg"
 repeat_region complement(16733..17040)
 /rpt_family="AluSx"
 repeat_region complement(17043..17355)
 /rpt_family="AluJb"
 repeat_region complement(17367..17655)
 /rpt_family="AluSg"
 repeat_region 17693..17995
 /rpt_family="AluY"
 repeat_region 18090..18234
 /rpt_family="MIR"

```

repeat_region complement(18716..19007)
repeat_region /rpt_family="AluY"
repeat_region 19113..19406
repeat_region /rpt_family="AluSx"
repeat_region 19482..19656
repeat_region /rpt_family="FRAM"
repeat_region 19716..20010
repeat_region /rpt_family="AluSg"
repeat_region complement(20084..20356)
repeat_region /rpt_family="AluSg"
repeat_region 20362..20459
repeat_region /rpt_family="AluSg/x"
repeat_region 20478..20675
repeat_region /rpt_family="AluSg/x"
repeat_region 20676..20976
repeat_region /rpt_family="AluSg"
repeat_region 21095..21320
repeat_region /rpt_family="AluSg/x"
repeat_region 21321..21397
repeat_region /rpt_family="AluSg/x"
repeat_region complement(22118..22380)
repeat_region /rpt_family="AluSg1"
repeat_region complement(22386..22495)
repeat_region /rpt_family="FLAM_C"
repeat_region 22521..22821
repeat_region /rpt_family="MIR"
repeat_region 22824..23121
repeat_region /rpt_family="AluSc"
repeat_region complement(23165..23249)
repeat_region /rpt_family="AluJo/FRAM"
repeat_region 23451..23490
repeat_region /rpt_family="MIR"
misc_feature 23623..24055
note="DDS similarity to overlapping ESTs:
(22630..24055) AA069831 zm67b07.s1 Stratagene
neuroepithelium (#937231) Homo sapiens cDNA clone 530677
3'; Score: 627 Identity: 395/449 (87%).-(23623..23966)
W48683 zc45h08.s1 Soares senescent fibroblasts NBHSF Homo
sapiens cDNA clone 325311 3';
Score: 609 Identity: 326/345 (94%).-Additional EST
matches:
H16570, H45028, R37825, Z40550, F02707, M79056"
gene complement(24035..33665)
/gene="FKBP38"
note="Human FK506-binding protein homologue encoding
gene"
CDS complement(join(24035..24121,24530..24661,25136..25213,
29467..29639,30082..30299,31237..31325,31417..31586,
33545..33665))
/gene="FKBP38"
/codon_start=1
/product="FK-506 binding protein homologue"
/protein_id="AAC28753.1"
/db_xref="GI:3395419"
/translation="MGQPEAEQPGALAREFLAAMEPEPAPAPAEWLDLNGSL
LRKTLVPGPSSRPVKGQVTVHLQTSLENTVQEEELFTLGDCDVIQALDLS
VPLMDVGETAMVTADSKYCYQPGRSYPPHAALEKLVKTAVDGPDLMLTQGER
VALANRREGNAHVQADVFIAANSYDIAIKAITSSAKVDMTFEEAQLLKLVKL
NLAASQLKDIHRAALRSCSLVLEHOPDNIKALFRKGVLAQOGEYSAPILRAL
KLEPNTKHAELSKLVKHAQRTETALYRKMLGNFSLPAKCPFGKANSIPKWL
FGATAVAGLVAISVVAARN"
/gene="FKBP38"
complement(25137..25214)
/note="FKBP38"
note="DDS similarity to HSU53083 Human clone 5/24 mRNA
sequence; (85..7); 97% identity."
repeat_region /rpt_family="AluSg/x"
Query Match 84.0%; Score 21; DB 9; Length 46213;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTTCAGTGAGCCGAGATAA 21

```

```

|||||
Db 9988 GGTTCAGTGAGCCGAGATAA 10008
|||||
RESULT 42
AB065918 49228 bp DNA linear PRI 23-JUL-2002
LOCUS cds, isolate:CBRC7TM_481.
DEFINITION Homo sapiens gene for seven transmembrane helix receptor, complete
ACCESSION AB065918
VERSION AB065918
KEYWORDS AB065918.1 GI:21925096
SOURCE Homo sapiens (isolate:CBRC7TM_481) DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Suwa.M., Sato.T., Okouchi.I., Arita.M., Futami.K., Matsumoto.S.,
Tsutsumi.S., Aburatani.H., Asai.K. and Akiyama.Y.
Genome-wide discovery and analysis of human seven transmembrane
helix receptor genes
Unpublished
2 (bases 1 to 49228)
Suwa.M.
Direct Submission
Submitted (11-JUL-2001) Makiko Suwa, Computational Biology Research
Center (CBRC), National Institute of Advanced Industrial Science
and Technology (AIST); 2-41-6 Aomi Koto-ku, Tokyo 135-0064, Japan
(E-mail:m-suwa@aist.go.jp, URL:http://www.cbrc.jp/,
Tel:81-3-3599-8080, Fax:81-3-3599-8081)
This sequence is a seven transmembrane helix receptor candidate
predicted from the whole human genome sequences using our automated
system that contains programs of gene
finding(GeneDecoder), sequence search, motif-domain assignment and
transmembrane helix prediction.
And the sequence is submitted by the collaborative project between
[Computational Biology Research Center (CBRC), National Institute
of Advanced Industrial Science and Technology (AIST)] and [Genome
Science Division, Research Center for Advanced Science and
Technology (RCAT), University of Tokyo].
FEATURES
source
1. 49228
/organism="Homo sapiens"
/isolate="CBRC7TM_481"
/db_xref="taxon:9606"
/chromosomes="19"
join(201..285,1211..1273,7128..7271,7879..8034,8158..8277,
12605..12751,14540..14680,14766..14912,17163..17255,
18348..18349,19419..19502,24377..24554,26975..27094,
30291..30490,32459..32629,35422..35616,37110..37345,
38889..38955,45730..45821,47989..48157,48290..48394,
48954..49028)
/codon_start=1
/evidence="not experimental"
/product="seven transmembrane helix receptor"
/protein_id="BAC06133.1"
/db_xref="GI:21929097"
/translation="MRSGDRHPGQGHGTFELLMKLAGRGCCVMHSGEHIRPTR
KPTNGNCRDSTLCPAYATCNTVDVSYCACQGLSSNGNHFKDPGVKRCNDIEC
SOSPQCCPNSSCKNLGSKYKCSLDGFSSTGNDVWVGKPNFSCDINECLTSVC
PHSDCVNMGVSQSCQVGFISRNCEDEVDRCADPRCEPHATCNTVGNYSQCN
PGFESSGHLSFQGLKASCEIDECETPCINTPTPGSYFCHPGFATNSQLN
FTDQGECDIDECDDPSTGPNISCTNAGLSYGGCIAGFHNPBGSKQDNFSQ
RVLFCKEDVLPDKNQIQOCEGTAVKPAVSVSCAONINNIERSVDLKVCEKNTIVS
LKNTEESFVPLKQISTWTKFTKEETSSLATVLESVSMFLASFWRKPSANTPAVRT
EYLDIESKVNKECSSENVTLDLVAKDKMKKGCSTIESESTETTGVAFFSVGVMS
VLNRPFDHQAPLTTSEIKLNMNSRVGGIMTGERKDFSDPIITLENVQPKQKE
RPICVSNSTVKGKRWTFPGCVILEASEYITICSNOMANLAVIMASGELTMDFFSLYI
ISHGIIISLVLCVIAITFLICRSIRNHNTHLHLCLVCLLTLFLAGHTDK
MGCALTAGLHYLFLACFFMLVEAVILFLMVRNLKVVNPESSNIMLHICARGYL
PMLVVISASVQPGYGMNRCMLNTGFIHSFLGPGVCTVIVINSLLTFTMLIQ
RLSSVNAEVLTKDPLTTFKAPQLFLGCSWLVGFIQGVAGVMAVLEFTINSLQ
GAFILHCLLNGVREYKRWITGKTPSSQSTSRILLSMPSASKTVSVIPATHE

```

BASE COUNT 12860 a 10969 c 10995 g 14404 t
 ORIGIN

Query Match 84.0%; Score 21; DB 9; Length 49228;
 Best Local Similarity 100.0%; Pred. No. 0.0038;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGATAA 21
 |||||
 Db 42476 GGTTCAGTGCAGCCGAGATAA 42496

RESULT 43

LOCUS AX015916

DEFINITION Sequence 19 from Patent WO9950284.

ACCESSION AX015916

VERSION AX015916.1 GI:10041659

KEYWORDS human.

SOURCE

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Aigner, T., Hess, J., Rosenthal, A., Rump, A. and Wirth, T.

TITLE Nucleic acid molecules which code proteins influencing bone development

JOURNAL PATENT: WO 9950284-A 19 07-OCT-1999;

ANDREAS (DE); WIRTH THOMAS (DE); ROSENTHAL ANDRE (DE); RUMP

FEATURES

source

1. .49999

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 10963 a 13723 c 13439 g 11854 t

ORIGIN

Query Match 84.0%; Score 21; DB 6; Length 49999;
 Best Local Similarity 100.0%; Pred. No. 0.0038;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGATAA 21
 |||||

Db 30972 GGTTCAGTGCAGCCGAGATAA 30992

RESULT 44

LOCUS AC068390

DEFINITION Homo sapiens clone RP11-303B11, LOW-PASS SEQUENCE SAMPLING.

ACCESSION AC068390

VERSION AC068390.1 GI:7677761

KEYWORDS HTG; HTGS-PHASE0.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.

TITLE Homo sapiens, clone RP11-303B11

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 52687)

AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
 Boguslavskiy, L., Bouckgalter, B., Brown, A., Burkett, G.,
 Campolano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
 Collymore, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S.,
 Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
 Grand-Pierre, N., Grant, G., Hagos, B., Hearford, A., Horton, L.,
 Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
 Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J.,

Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
 Meidrum, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
 Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
 Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
 Stange-rhmann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission

TITLE

JOURNAL

COMMENT

Submitted (02-MAY-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WISR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: LI0258

Center clone name: 303_B_11

* NOTE: This record contains 68 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1 691: contig of 691 bp in length
 692 791: gap of 100 bp
 792 1466: contig of 675 bp in length
 1467 1566: gap of 100 bp
 1567 2246: contig of 680 bp in length
 2247 2346: gap of 100 bp
 2347 3025: contig of 679 bp in length
 3026 3125: gap of 100 bp
 3126 3810: contig of 685 bp in length
 3811 3910: gap of 100 bp
 3911 4604: contig of 694 bp in length
 4605 4704: gap of 100 bp
 4705 5311: contig of 607 bp in length
 5312 5411: gap of 100 bp
 5412 6070: contig of 659 bp in length
 6071 6170: gap of 100 bp
 6171 6850: contig of 680 bp in length
 6851 6950: gap of 100 bp
 6951 7637: contig of 687 bp in length
 7638 7737: gap of 100 bp
 7738 8402: contig of 665 bp in length
 8403 8502: gap of 100 bp
 8503 9182: contig of 680 bp in length
 9183 9282: gap of 100 bp
 9283 9967: contig of 685 bp in length
 9968 10067: gap of 100 bp
 10068 10763: contig of 696 bp in length
 10764 10863: gap of 100 bp
 10864 11558: contig of 695 bp in length
 11559 11658: gap of 100 bp
 11659 12281: contig of 623 bp in length
 12282 12381: gap of 100 bp
 12382 13074: contig of 693 bp in length
 13075 13174: gap of 100 bp
 13175 13852: contig of 678 bp in length
 13853 13952: gap of 100 bp

```
* 13953 14630: contig of 678 bp in length
* 14631 14730: gap of 100 bp
* 14731 15394: contig of 664 bp in length
* 15395 15494: gap of 100 bp
* 15495 16163: contig of 675 bp in length
* 16170 16269: gap of 100 bp
* 16270 16954: contig of 685 bp in length
* 16955 17054: gap of 100 bp
* 17055 17747: contig of 693 bp in length
* 17748 17847: gap of 100 bp
* 17848 18497: contig of 650 bp in length
* 18498 18597: gap of 100 bp
* 18598 19235: contig of 638 bp in length
* 19236 19335: gap of 100 bp
* 19336 20028: contig of 693 bp in length
* 20029 20128: gap of 100 bp
* 20129 20797: contig of 669 bp in length
* 20798 20897: gap of 100 bp
* 20898 21594: contig of 697 bp in length
* 21595 21694: gap of 100 bp
* 21695 22358: contig of 664 bp in length
* 22359 22458: gap of 100 bp
* 22459 23136: contig of 678 bp in length
* 23137 23236: gap of 100 bp
* 23237 23899: contig of 663 bp in length
* 23900 23999: gap of 100 bp
* 24000 24663: contig of 664 bp in length
* 24664 24763: gap of 100 bp
* 24764 25447: contig of 684 bp in length
* 25448 25547: gap of 100 bp
* 25548 26238: contig of 691 bp in length
* 26239 26338: gap of 100 bp
* 26339 27032: contig of 694 bp in length
* 27033 27132: gap of 100 bp
* 27133 27817: contig of 685 bp in length
* 27818 27917: gap of 100 bp
* 27918 28601: contig of 684 bp in length
* 28602 28701: gap of 100 bp
* 28702 29381: contig of 680 bp in length
* 29382 29481: gap of 100 bp
* 29482 30162: contig of 681 bp in length
* 30163 30262: gap of 100 bp
* 30263 30881: contig of 619 bp in length
* 30882 30981: gap of 100 bp
* 30982 31666: contig of 685 bp in length
* 31667 31766: gap of 100 bp
* 31767 32421: contig of 655 bp in length
* 32422 32521: gap of 100 bp
* 32522 33207: contig of 686 bp in length
* 33208 33307: gap of 100 bp
* 33308 33983: contig of 676 bp in length
* 33984 34083: gap of 100 bp
* 34084 34762: contig of 679 bp in length
* 34763 34862: gap of 100 bp
* 34863 35542: contig of 680 bp in length
* 35543 35642: gap of 100 bp
* 35643 36333: contig of 691 bp in length
* 36334 36433: gap of 100 bp
* 36434 37119: contig of 686 bp in length
* 37120 37219: gap of 100 bp
* 37220 37909: contig of 690 bp in length
* 37910 38009: gap of 100 bp
* 38010 38654: contig of 645 bp in length
* 38655 38754: gap of 100 bp
* 38755 39446: contig of 692 bp in length
* 39447 39546: gap of 100 bp
* 39547 40222: contig of 676 bp in length
* 40223 40322: gap of 100 bp
* 40323 41011: contig of 689 bp in length
* 41012 41111: gap of 100 bp
* 41112 41788: contig of 677 bp in length
* 41789 41888: gap of 100 bp
* 41889 42574: contig of 686 bp in length
```

```
* 42575 42674: gap of 100 bp
* 42675 43366: contig of 692 bp in length
* 43367 43466: gap of 100 bp
* 43467 44142: contig of 676 bp in length
* 44143 44242: gap of 100 bp
* 44243 44924: contig of 682 bp in length
* 44925 45024: gap of 100 bp
* 45025 45705: contig of 681 bp in length
* 45706 45805: gap of 100 bp
* 45806 46461: contig of 656 bp in length
* 46462 46561: gap of 100 bp
* 46562 47224: contig of 663 bp in length
* 47225 47324: gap of 100 bp
* 47325 47996: contig of 672 bp in length
* 47997 48096: gap of 100 bp
* 48097 48777: contig of 681 bp in length
* 48778 48877: gap of 100 bp
* 48878 49553: contig of 676 bp in length
* 49554 49653: gap of 100 bp
* 49654 50323: contig of 670 bp in length
* 50324 50423: gap of 100 bp
* 50424 51112: contig of 689 bp in length
* 51113 51212: gap of 100 bp
* 51213 51901: contig of 689 bp in length
* 51902 52001: gap of 100 bp
* 52002 52687: contig of 686 bp in length.

FEATURES             Location/Qualifiers
     source            1..52687
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
```

```
Query Match          84.0%; Score 21; DB 2; Length 52687;
Best Local Similarity 100.0%; Pred. No. 0.0038;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGTGCGAGTGCAGCCGAGATAA 21
|||||
Db 26192 GGTGCGAGTGCAGCCGAGATAA 26212
```

RESULT 45

AC091585/c

```
LOCUS               59305 bp   DNA   linear   HTG 28-NOV-2001
DEFINITION          Homo sapiens chromosome 18 clone RP11-46619 map 18, LOW-PASS
                     SEQUENCE SAMPLING.
ACCESSION            AC091585
VERSION              AC091585.2 GI:17047205
KEYWORDS             HTG; HTGS_PHASE0.
SOURCE               Homo sapiens.
ORGANISM             Homo sapiens
```

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

Bairren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,

Barra,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B., Brown,A.,

Cammarata,J., Campopiano,A., Chang,J., Choepel,Y., Colangelo,M.,

Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K.,

Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D.,

Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,

Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W.,

Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRocque,K.,

Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,

McClean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M.,

McEwan,T., McKernan,K., McPheeters,R., Meldrum,J., Meneus,L.,

Milova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,

Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,

Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,

Retta,R., Rieback,M., Riley,K., Rise,C., Rogov,P., Roman,J.,

Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,

Severy,P., Sougnez,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (07-MAY-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Nov 22, 2001 this sequence version Replaced gi:13959202.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L13172

Center clone name: 466_I_9

*** NOTE: This record contains 72 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 746: contig of 746 bp in length
747 846: gap of 100 bp
847 1562: contig of 716 bp in length
1563 1662: gap of 100 bp
1663 2397: contig of 735 bp in length
2398 2497: gap of 100 bp
2498 3220: contig of 723 bp in length
3221 3320: gap of 100 bp
3321 4049: contig of 729 bp in length
4050 4149: gap of 100 bp
4150 4883: contig of 734 bp in length
4884 4983: gap of 100 bp
4984 5666: contig of 683 bp in length
5667 5766: gap of 100 bp
5767 6499: contig of 733 bp in length
6500 6599: gap of 100 bp
6600 7301: contig of 702 bp in length
7302 7401: gap of 100 bp
7402 8118: contig of 717 bp in length
8119 8218: gap of 100 bp
8219 8946: contig of 728 bp in length
8947 9046: gap of 100 bp
9047 9793: contig of 747 bp in length
9794 9893: gap of 100 bp
9894 10624: contig of 731 bp in length
10625 10724: gap of 100 bp
10725 11459: contig of 735 bp in length
11460 11559: gap of 100 bp
11560 12284: contig of 725 bp in length
12285 12384: gap of 100 bp
12385 13095: contig of 711 bp in length
13096 13195: gap of 100 bp
13196 13922: contig of 727 bp in length
13923 14022: gap of 100 bp
14023 14748: contig of 726 bp in length
14749 14848: gap of 100 bp
14849 15564: contig of 716 bp in length
15565 15664: gap of 100 bp
15665 16351: contig of 687 bp in length
16352 16451: gap of 100 bp
16452 17171: contig of 720 bp in length

17172 17271: gap of 100 bp
17272 17993: contig of 722 bp in length
17994 18093: gap of 100 bp
18094 18825: contig of 732 bp in length
18826 18925: gap of 100 bp
18926 19649: contig of 724 bp in length
19650 19749: gap of 100 bp
19750 20471: contig of 722 bp in length
20472 20571: gap of 100 bp
20572 21301: contig of 730 bp in length
21302 21401: gap of 100 bp
21402 22120: contig of 719 bp in length
22121 22220: gap of 100 bp
22221 22959: contig of 739 bp in length
22960 23059: gap of 100 bp
23060 23810: contig of 751 bp in length
23811 23910: gap of 100 bp
23911 24657: contig of 747 bp in length
24658 24757: gap of 100 bp
24758 25478: contig of 721 bp in length
25479 25578: gap of 100 bp
25579 26313: contig of 735 bp in length
26314 26413: gap of 100 bp
26414 27122: contig of 709 bp in length
27123 27222: gap of 100 bp
27223 27969: contig of 747 bp in length
27970 28069: gap of 100 bp
28070 28803: contig of 734 bp in length
28804 28903: gap of 100 bp
28904 29641: contig of 738 bp in length
29642 29741: gap of 100 bp
29742 30465: contig of 724 bp in length
30466 30565: gap of 100 bp
30566 31301: contig of 736 bp in length
31302 31401: gap of 100 bp
31402 32127: contig of 726 bp in length
32128 32227: gap of 100 bp
32228 32963: contig of 736 bp in length
32964 33063: gap of 100 bp
33064 33775: contig of 712 bp in length
33776 33875: gap of 100 bp
33876 34597: contig of 722 bp in length
34598 34697: gap of 100 bp
34698 35425: contig of 728 bp in length
35426 35525: gap of 100 bp
35526 36230: contig of 705 bp in length
36231 36330: gap of 100 bp
36331 37058: contig of 728 bp in length
37059 37158: gap of 100 bp
37159 37879: contig of 721 bp in length
37880 37979: gap of 100 bp
37980 38701: contig of 722 bp in length
38702 38801: gap of 100 bp
38802 39526: contig of 725 bp in length
39527 39626: gap of 100 bp
39627 40336: contig of 710 bp in length
40337 40436: gap of 100 bp
40437 41169: contig of 733 bp in length
41170 41269: gap of 100 bp
41270 41999: contig of 730 bp in length
42000 42099: gap of 100 bp
42100 42837: contig of 738 bp in length
42838 42937: gap of 100 bp
42938 43645: contig of 708 bp in length
43646 43745: gap of 100 bp
43746 44464: contig of 719 bp in length
44465 44564: gap of 100 bp
44565 45302: contig of 738 bp in length
45303 45402: gap of 100 bp
45403 46109: contig of 707 bp in length
46110 46209: gap of 100 bp
46210 46919: contig of 710 bp in length
46920 47019: gap of 100 bp

```

* 47020 47753: contig of 734 bp in length
* 47754 47853: gap of 100 bp
* 47854 48572: contig of 719 bp in length
* 48573 48672: gap of 100 bp
* 48673 49391: contig of 719 bp in length
* 49392 49491: gap of 100 bp
* 49492 50210: contig of 719 bp in length
* 50211 50310: gap of 100 bp
* 50311 51042: contig of 732 bp in length
* 51043 51142: gap of 100 bp
* 51143 51873: contig of 731 bp in length
* 51874 51973: gap of 100 bp
* 51974 52713: contig of 740 bp in length
* 52714 52813: gap of 100 bp
* 52814 53554: contig of 741 bp in length
* 53555 53654: gap of 100 bp
* 53655 54379: contig of 725 bp in length
* 54380 54479: gap of 100 bp
* 54480 55211: contig of 732 bp in length
* 55212 55311: gap of 100 bp
* 55312 56048: contig of 737 bp in length
* 56049 56148: gap of 100 bp
* 56149 56876: contig of 728 bp in length

Query Match      84.0%; Score 21; DB 2; Length 59305;
Best Local Similarity 100.0%; Pred. No. 0.0038;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGTTCAGTGAGCCGAGATAA 21
    |||||
DB 55482 GGTTCAGTGAGCCGAGATAA 55462

```

Search completed: November 24, 2002, 14:17:41
Job time : 2087.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2002, 13:27:24 ; Search time 1549.5 Seconds
(without alignments)
261.302 Million cell updates/sec

Title: US-09-784-423-125

Perfect score: 25

Sequence: 1 TGTGCCAGACACGAAATTACAG 25

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	18	72.0	838	12	BG414651
c 2	18	72.0	977	17	CNS046T0
c 3	17	68.0	287	9	AV296392
4	17	68.0	522	9	AA865753
5	17	68.0	642	14	BQ269298
6	17	68.0	821	12	BE877455

c 7	17	68.0	857	13	BM401231
c 8	17	68.0	961	14	BQ711606
c 9	16	64.0	229	14	BQ280096
c 10	16	64.0	246	10	AV357770
c 11	16	64.0	311	10	B8534005
c 12	16	64.0	320	9	AV219319
c 13	16	64.0	394	12	BG158653
c 14	16	64.0	395	14	AZ282651
c 15	16	64.0	396	17	AZ282651
c 16	16	64.0	397	14	T79727
c 17	16	64.0	397	14	T79727
c 18	16	64.0	408	17	BH878510
c 19	16	64.0	447	17	BH876286
c 20	16	64.0	449	9	AA126284
c 21	16	64.0	468	9	AI325084
c 22	16	64.0	471	17	BH879621
c 23	16	64.0	489	17	B98791
c 24	16	64.0	505	12	BF610439
c 25	16	64.0	533	14	BM725916
c 26	16	64.0	543	14	BQ701598
c 27	16	64.0	551	17	BH876388
c 28	16	64.0	557	12	BG039852
c 29	16	64.0	574	10	AW644504
c 30	16	64.0	581	17	BH876387
c 31	16	64.0	586	17	BH876535
c 32	16	64.0	601	10	AW043028
c 33	16	64.0	601	10	AW043050
c 34	16	64.0	601	10	AW043091
c 35	16	64.0	608	17	AE009748
c 36	16	64.0	613	10	BE000467
c 37	16	64.0	628	12	BG033085
c 38	16	64.0	636	12	BF511713
c 39	16	64.0	663	14	BM678024
c 40	16	64.0	667	17	AG133215
c 41	16	64.0	686	17	BH882704
c 42	16	64.0	698	10	AW340752
c 43	16	64.0	752	17	AQ745078
c 44	16	64.0	806	12	BG681831
c 45	16	64.0	825	12	BF978854

ALIGNMENTS

RESULT 1

BG414651/c

LOCUS

DEFINITION

HVSMEK0003D14f Hordeum vulgare testa/pericarp EST library

HYCDNA0013 (normal) Hordeum vulgare cDNA clone HVSMEK0003D14f, mRNA

sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BG414651 838 bp mRNA linear EST 23-OCT-2001
HVSMEK0003D14f Hordeum vulgare testa/pericarp EST library
HYCDNA0013 (normal) Hordeum vulgare cDNA clone HVSMEK0003D14f, mRNA
sequence.
ACCESSION BG414651.2 GI:16328533
VERSION BG414651
KEYWORDS EST.
SOURCE Hordeum vulgare
ORGANISM Hordeum vulgare
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS ; Triticeae; Hordeum.
; Triticeae; Hordeum.
Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Kannangara,G., von
Wetstein,D., Ahtunov,E., Chin,A., Choi,D.W., Fenton,R.D., Kianlian
P., Otto,C., Simons,K., Zhang,D., Begum,D., Frisch,D., Yu,Y.,
Henry,D., Palmer,M., Rambo,T., Simmons,J., Oates,R. and Main,D.
Development of a genetically and physically anchored EST resource
for barley genomics: Morex testa/pericarp cDNA library
Unpublished (2001)
On Mar 13, 2001 this sequence version replaced gi:13320202.
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293

Email: rwing@clmson.edu
Total hg bases = 117
Seq primer: AATTACCTCCTAATGAGG
High quality sequence stop: 502.
Location/Qualifiers
1. .838
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSX0003D14f"
/clone_lib="Hordeum vulgare testa/pericarp EST library
HvCDNA0013 (normal)"
/tissue_type="testa/pericarp"
/lab_host="TJC121"
/note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI;
Plants were raised from seeds in a Controlled Environments
growth chamber maintained in continuous light at 18°C, and
testa and pericarp were dissected from developing kernels
at Washington State University, Pullman, WA (Rannangara,
von Wettstein). Total RNA was prepared, poly(A) RNA was
purified, one cDNA library was made, and 1 million pfu
were in vivo excised to give pBluescript SK(-) cDNA
phagemids in the TJ Close lab at the University of
California, Riverside (Akhunov, Chin, Choi, Close, Fenton,
Kilian, Otto, Simons, Zhang). Phagemids were plated and
picked at the Clemson University Genomics Institute (CUGI)
(Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA
preparations, DNA sequencing and sequence analysis were
performed at CUGI (Wing, Yu, Frisch, Henry, Simons, Oates
Rambo, Main). The sequence has been trimmed to remove
vector sequence and contains a minimum of 100 bases of
phred value 20 or above. For more details on library
preparation and sequence analysis see
<http://www.genome.clemson.edu/projects/barley>. To order
this clone see <http://www.genome.clemson.edu/orders> Also
see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)"

BASE COUNT 208 a 196 c 178 g 236 t
ORIGIN

Query Match 72.0%; Score 18; DB 12; Length 838;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GCCAGGACCGAATTT 21
|||||
Db 620 GCCAGGACCGAATTT 603

RESULT 2
CNS046T0/c
LOCUS
DEFINITION
Tetraodon nigroviridis genome survey sequence pUC-ori end of clone
086422 of library G from Tetraodon nigroviridis, genomic survey
sequence.
AL277101
Accession
AL277101.1 GI:8011304
VERSION
GSS: genome survey sequence.
KEYWORDS
Tetraodon nigroviridis.
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1 (bases 1 to 977)
Roest-Crolius.H., Jaillon.O., Dasilva.C., Bouneau.L., Fisher.C.,
Bernot.A., Fizames.C., Wincker.P., Brottier.P., Quetier.F.,
Saurin.W. and Weissenbach.J.
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished

REFERENCE 2 (bases 1 to 977)
AUTHORS
Roest-Crolius.H., Jaillon.O., Dasilva.C., Fizames.C., Fisher.C.,
Bouneau.L., Billault.A., Quetier.F., Saurin.W., Bernot.A. and
Weissenbach.J.
TITLE
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL
Unpublished
REFERENCE 3 (bases 1 to 977)
AUTHORS
Genoscope.
TITLE
Direct Submission
JOURNAL
Submitted (12-APR-2000)
COMMENT
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
<http://www.genoscope.cns.fr/tetraodon>.

FEATURES
source
1. .977
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="086A23"
/clone_lib="G"
/note="Genoscope sequence ID : C0B08069A11SPI-end :
PUC-ori"

BASE COUNT 305 a 185 c 161 g 306 t 20 others
ORIGIN

Query Match 72.0%; Score 18; DB 17; Length 977;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AGGAACCGAATTTTACA 24
|||||
Db 177 AGGAACCGAATTTTACA 160

RESULT 3
AV296392/c
LOCUS
DEFINITION
AV296392 RIKEN full-length enriched, 8 days embryo Mus musculus
cDNA clone 5730441H16 3' similar to X15013 Rat mRNA for ribosomal
protein L7a, mRNA sequence.
AV296392 287 bp mRNA linear EST 10-NOV-1999
AV296392.1 GI:6328409
VERSION
EST.
KEYWORDS
house mouse.
SOURCE
ORGANISM
Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 287)
Konno.H., Aizawa.K., Akahira.S., Akiyama.J., Carninci.P., Endo.T.,
Fukuda.S., Fukunishi.Y., Hara.A., Hayatsu.N., Hirozane.T., Hori.F.,
Ishii.Y., Ishikawa.T., Itoh.M., Izawa.M., Kadota.K., Kagawa.I., Kai
.C., Kawai.J., Kikuchi.N., Kojima.Y., Koya.S., Kusakabe.M.,
Matsuyama.T., Miki.R., Mizuno.Y., Nakamura.M., Oda.H., Okazaki.Y.,
Owa.C., Ozawa.Y., Saifu.H., Sano.M., Sato.K., Shibata.K., Shibata
.Y., Shigenoto.Y., Shiraki.T., Sogabe.Y., Sugahara.Y., Suzuki.H.,
Suzuki.H., Takahashi.P., Tateo.M., Tominaga.N., Tsunoda.Y.,
Watahiki.A., Watanabe.S., Yamamura.T., Yasunishi.A., Yokota.T.,
Yoshiki.A., Yoshino.M., Muramatsu.M. and Hayashizaki.Y.
RIKEN Mouse ESTs (Konno.H., et al. 1999)
Unpublished (1999)
Contact: Yoshinide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.riken.go.jp,
URL:<http://genome.gsc.riken.go.jp/>
Sasaki.N., Izawa.M., Watahiki.M., Ozawa.K., Tanaka.T., Yoneda.Y.,
Matsura.S., Carninci.P., Muramatsu.M., Okazaki.Y. and Hayashizaki
.Y.

(hinoue@im.wustl.edu)
 Seq primer: -40RP from Gibco
 High quality sequence stop: 481.
 Location/Qualifiers
 1. .642

FEATURES

source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:388094"
 /clone_lib="HR85 islet"
 /tissue_type="Purified pancreatic islet"
 /lab_host="DH10B"

/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1: NotI; Site_2: XhoI; cDNA made by oligo-dT priming. Size: selected on agarose gel. Average insert size -1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permutt Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."

BASE COUNT
 ORIGIN

166 a 179 c 190 g 107 t

Query Match 68.0%; Score 17; DB 14; Length 642;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 TCCAGGACCAAGAAAT 19
 Db 409 TCCAGGACCAAGAAAT 425

RESULT 6

BE877455

LOCUS

DEFINITION 601485493F1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:388094 5',
 mRNA sequence.

ACCESSION

BE877455

VERSION

BE877455.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 821)

AUTHORS

NIH-MGC http://mhc.nci.nih.gov/.

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCRD/DP/Gazdar
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM9667 row: i column: 23
 High quality sequence stop: 715.

FEATURES

source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:388094"
 /clone_lib="NIH_MGC_69"
 /tissue_type="large cell carcinoma, undifferentiated"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pCMV-Sport6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.1 kb. Library constructed by Life
 Technologies."

BASE COUNT
 ORIGIN

228 a 270 c 115 g 208 t

Query Match 68.0%; Score 17; DB 12; Length 821;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AGGAACCAAGAAATTTAC 23
 Db 713 AGGAACCAAGAAATTTAC 729

RESULT 7

BM401231/c

LOCUS

DEFINITION 5009-0-84-E04.t.1 Chilcoat/Turkewitz cDNA (large fraction)
 Tetrahymena thermophila cDNA, mRNA sequence.

ACCESSION

BM401231

VERSION

BM401231.1

KEYWORDS

EST.

SOURCE

Tetrahymena thermophila.

ORGANISM

Tetrahymena thermophila.

REFERENCE

1 (bases 1 to 857)

AUTHORS

Turkewitz, A.P., Karrer, K.M., Jahn, C., Orlas, E., Kirk, K.E., Frankel
 J., and Klobutcher, L.

JOURNAL

Unpublished (2002)

COMMENT

EST from Tetrahymena thermophila, strain CU428.1, growing cells
 Contact: Turkewitz AP
 Molecular Genetics and Cell Biology
 University of Chicago
 920 E. 58th Street, Chicago, IL 60637, USA
 Tel: 773 702 4374
 Fax: 773 702 3172
 Email: apturkew@midway.uchicago.edu
 Seq primer: T3.

FEATURES

source

1. .857

Location/Qualifiers

/organism="Tetrahymena thermophila"

/strain="CU428.1"

/db_xref="taxon:5911"

/clone_lib="Chilcoat/Turkewitz cDNA (large fraction)"

/note="Vector: Bluescript2 SK+; Details on library
 preparation can be found in Chilcoat and Turkewitz (2001)
 Proc. Natl. Acad. Sci USA, 98: 8709-8713."

BASE COUNT

196 a 165 c 213 g 282 t

ORIGIN

1 others

Query Match 68.0%; Score 17; DB 13; Length 857;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGAACCAAGAA 17
 Db 408 TGTGCCAGGAACCAAGAA 392

RESULT 8

BQ711606

LOCUS

DEFINITION AGENCOURT_8501294 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6302959
 5', mRNA sequence.

ACCESSION

BQ711606

VERSION

BQ711606.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 961)

AUTHORS

NIH-MGC http://mhc.nci.nih.gov/.

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-re@mail.nih.gov
 Tissue Procurement: Dr. Mark Watson
 cDNA Library Preparation: Rubin Laboratory
 DNA Sequencing Arrayed By: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution by: Agencourt Bioscience Corporation
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LICM2521 row: e column: 08
 High quality sequence start: 8
 High quality sequence stop: 542.
 Location/Qualifiers
 1..961

FEATURES

source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6302959"
 /clone_lib="NIH_MGC_113"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: spleen; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGACAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH MGC Library."

BASE COUNT 237 a 260 c 227 g 236 t 1 others
 ORIGIN

Query Match 68.0%; Score 17; DB 14; Length 961;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 CCAGGACCAAGATTT 21
 Db 829 CCAGGACCAAGATTT 845
 RESULT 9
 BQ280096/c
 LOCUS
 DEFINITION BQ280096 229 bp mRNA linear EST 08-MAY-2002
 Schmidt lab Zea mays cDNA, mRNA sequence.
 ACCESSION BQ280096
 VERSION BQ280096.1 GI:20512910
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 229)
 Walbot,V.
 Maize ESTs from various cDNA libraries sequenced at Stanford
 University
 Unpublished (1999)
 Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 1091034 row: F column: 06.
 Location/Qualifiers
 1..229
 /organism="Zea mays"
 /cultivar="OH43"
 /db_xref="taxon:4577"
 /clone_lib="1091 - Immature ear with common ESTs screened
 by Schmidt lab"
 /tissue_type="Inflorescence meristem - floral organ
 primordia"

FEATURES

source

Location/Qualifiers
 1..246
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="7420408P18"
 /clone_lib="RIKEN full-length enriched, in vitro

/dev_stage="0.5 cm to 2 cm"
 /lab_host="Stratagene XL0LR"
 Note="Organ: Immature ear; Vector: pAD-GAL4; Site_1:
 EcoRI; Site_2: XhoI; RNA from library 606 was filtered for
 common ESTs found in 606."
 BASE COUNT 80 a 32 c 38 g 79 t
 ORIGIN

Query Match 64.0%; Score 16; DB 14; Length 229;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGTGCCAGGAACCAGA 16
 Db 43 TGTGCCAGGAACCAGA 28
 RESULT 10
 AV357770/c
 LOCUS
 DEFINITION AV357770 RIKEN full-length enriched, in vitro fertilized eggs Mus
 musculus cDNA clone 7420408P18 3', mRNA sequence.
 ACCESSION AV357770
 VERSION AV357770.1 GI:6404772
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 246)
 Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T.,
 Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F.,
 Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai
 ,C., Kawai,J., Kikuchi,N., Koijima,Y., Koya,S., Kusakabe,M., Kai
 Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,
 Owa,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibata,K., Shibata
 ,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H.,
 Suzuki,H., Takahashi,F., Tateo,M., Tominaga,N., Tsunoda,Y.,
 Wathiki,A., Watanabe,S., Yamamura,T., Yasunishi,A., Yokota,T.,
 Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
 RIKEN Mouse ESTs (Konno,H., et al. 1999)
 Unpublished (1999)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsr.riken.go.jp,
 URL: <http://genome.gsc.riken.go.jp/>
 Sasaki,N., Izawa,M., Wathiki,M., Ozawa,K., Tanaka,T., Yoneda,Y.,
 Matsuura,S., Carninci,P., Muramatsu,M., Okazaki,Y. and Hayashizaki
 ,Y.
 Transcriptional sequencing: A method for DNA sequencing using RNA
 polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
 Itoh,M., Kitsumai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
 Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
 ,Y. and Hayashizaki,Y.
 Automated filtration-based high-throughput plasmid preparation
 system. Genome Res. 9 (5), 463-470 (1999)
 Carninci,P. and Hayashizaki,Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
 19-44 (1999)
 Please visit our web site (<http://genome.ritc.riken.go.jp>) for
 further details.

FEATURES

source

Location/Qualifiers
 1..246
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="7420408P18"
 /clone_lib="RIKEN full-length enriched, in vitro

```

fertilized eggs"
/sex="female"
/tissue_type="in vitro fertilized eggs"
/dev_stage="egg"
/lab_host="DH10B"
/note="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGATTCTCGATTAAATTAATTCCTCCCCCCCC 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I. Cloning sites, 5' end: Sali; 3' end: BamHI"
BASE COUNT      69 a   33 c   42 g   101 t   1 others
ORIGIN
Query Match      64.0%; Score 16; DB 10; Length 246;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AGGAACCAAGATTTA 22
|||||
Db 234 AGGAACCAAGATTTA 219

RESULT 11
BB534005
LOCUS
DEFINITION
BB534005 311 bp mRNA linear EST 29-JUL-2000
musculus cDNA clone E030032P21 3' similar to I28901 Rat
transcription factor IIIC alpha-subunit mRNA, mRNA sequence.
ACCESSION
BB534005.1 GI:9585934
VERSION
KEYWORDS
SOURCE
EST.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 311)
Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci
,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.,
Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,
Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata
,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y.,
Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tomimaga,N., Toya
,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I.,
Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiaki,A., Yoshino
,M., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Konno,H., et al.)
Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki
,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermotabilization and thermoactivation of thermolabile enzymes by

```

```

trehalose and its application for the synthesis of full length
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsuina,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for
further details.
Location/Qualifiers
1..311
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="E030032P21"
/clone_lib="RIKEN full-length enriched, 0 day neonate
lung"
/tissue_type="lung"
/dev_stage="0 day neonate"
/lab_host="DH10B"
/note="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGGCGCGCACTCGAGTTTCTTTTCTTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGATTCTCGATTAAATTAATTCCTCCCCCCCC 3']. cDNA
was cleaved with BamHI and XhoI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLC I."
BASE COUNT      103 a   89 c   68 g   51 t
ORIGIN
Query Match      64.0%; Score 16; DB 10; Length 311;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

FEATURES

source

```

Location/Qualifiers
1..311
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="E030032P21"
/clone_lib="RIKEN full-length enriched, 0 day neonate
lung"
/tissue_type="lung"
/dev_stage="0 day neonate"
/lab_host="DH10B"
/note="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGGCGCGCACTCGAGTTTCTTTTCTTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGATTCTCGATTAAATTAATTCCTCCCCCCCC 3']. cDNA
was cleaved with BamHI and XhoI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLC I."
BASE COUNT      103 a   89 c   68 g   51 t
ORIGIN
Query Match      64.0%; Score 16; DB 10; Length 311;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 3 TCACAGACACAGAAA 18

|||||

Db 203 TCACAGACACAGAAA 218

AV219319

AV219319/c

LOCUS

DEFINITION

musculus cDNA clone 3010087J17 3' similar to U19940 Fragaria x

ananassa putative 40S ribosomal protein s12 mRNA, mRNA sequence.

AV219319

ACCESSION

VERSION

KEYWORDS

SOURCE

EST.

ORGANISM

house mouse

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 320)

Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T.,

Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F.,

Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,T., Kai

,C., Kawai,J., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M.,

Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,

Owa,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibata,K., Shibata

,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H.,

Suzuki,H., Takahashi,F., Tateno,M., Tomimaga,N., Tsunoda,Y.,

Watahiki,A., Watanabe,S., Yamamura,T., Yasunishi,A., Yokota,T.,

Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

FEATURES
source

Location/Qualifiers
1. 395
/organism="Homo sapiens"
/db_xref="GDB:475349"
/db_xref="taxon:9606"
/clone="IMAGE:122804"
/clone_lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia) with a modified polylinker; Site.1: Pac I; Site.2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' ACTCGAGAGATTTATTAAGATCTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Patima Bonaldo."

BASE COUNT 110 a 89 c 92 g 99 t 5 others
ORIGIN

Query Match 64.0%; Score 16; DB 14; Length 395;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TCCAGGAGACAGAAA 18
|||||
Db 50 TCCAGGAGACAGAAA 65

RESULT 15

AZ282651/c
LOCUS AZ282651 396 bp DNA linear GSS 27-JUL-2000
DEFINITION RPCI-23-152D24.TJ RPCI-23 Mus musculus genomic clone RPCI-23-152D24
, DNA sequence.

ACCESSION AZ282651
VERSION AZ282651.1 GI:9524437
KEYWORDS GSS.
SOURCE house mouse.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 396)
Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
and Fraser,C.M.

Mouse BAC End Sequences from Library RPCI-23

TITLE

JOURNAL

COMMENT

Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics (inforesgen.com). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 152 row: D column: 24
Seq primer: SP6
Class: BAC ends.

FEATURES
source

Location/Qualifiers
1. 396
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-152D24"
/clone_lib="RPCI-23"
/sex="Female"

/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site.1:
EcoRI; Site.2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBAC3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 70 a 126 c 61 g 139 t
ORIGIN

Query Match 64.0%; Score 16; DB 17; Length 396;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TCCAGGAGACAGAAA 18
|||||
Db 169 TCCAGGAGACAGAAA 154

RESULT 16

T77962
LOCUS T77962 397 bp mRNA linear EST 07-MAR-1995
DEFINITION YC97D11.r1 Soares infant brain 1N1B Homo sapiens cDNA clone
IMAGE:24089 5', mRNA sequence.

ACCESSION T77962
VERSION T77962.1 GI:696471
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 397)
Hillier,L., Clark,N., Dubouque,T., Elliston,K., Hawkins,M., Holman
M., Bultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,B., Waterston
R., Williamson,A., Wohlmann,P. and Wilson,R.

The WashU-Merck EST Project

TITLE

JOURNAL

COMMENT

Contact: Wilton RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1685

High quality sequence stops: 310 Source: IMAGE Consortium, LML
This clone is available royalty-free through LML; contact the
IMAGE Consortium (infoimage.llnl.gov) for further information.
Insert Length: 1685 Std Error: 0.00
Seq primer: M13RP1

High quality sequence stop: 310.

FEATURES

Location/Qualifiers
1. 397

/organism="Homo sapiens"
/db_xref="GDB:396436"
/db_xref="taxon:9606"
/clone="IMAGE:24089"
/clone_lib="Soares infant brain 1N1B"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: whole brain; Vector: Lamiid BA; Site.1: Not
I; Site.2: Hind III; 1st strand cDNA was primed with a Not
I - oligo(dT) primer [5'
AACTGAGAGATTCGGCCGACAGAAATTTTTTTTTTTT 3'];
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the Lamiid BA vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M.Patima Bonaldo."

BASE COUNT 85 a 115 c 92 g 102 t 3 others
ORIGIN

```

Query Match          64.0%; Score 16; DB 14; Length 397;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTGCCAGGACCAAGAA 17
|||||
Db 193 GTGCCAGGACCAAGAA 208

RESULT 17
BH878510          408 bp      DNA      linear      GSS 05-AUG-2002
LOCUS             hs82e08.bl WGS-ZmaysF (JM107 adapted methyl filtered) Zea mays
DEFINITION        genomic clone hs82e08 5', DNA sequence.
ACCESSION         BH878510
VERSION           BH878510.1 GI:22114407
KEYWORDS          GSS.
SOURCE            Zea mays.
ORGANISM          Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 408)
Rabinowicz,P.D., O'Shaughnessy,A.L., Balijs,V., Dedhia,N.,
Katzenburger,F., King,L., Miller,B., Muller,S., Nascimento,L.,
Zutavern,T., McCombie,W.R. and Martienssen,R.A.
Genomic shotgun sequences from Zea mays (methyl-filtered)
Unpublished (2002)
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: hs82 row: e column: 08
Seq primer: -21M13UnivFwd
Class: Shotgun
High quality sequence stop: 408.
Location/Qualifiers
1. 408
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="hs82e08"
/clone_host="JM107 or DH5a"
/label="Organ: immature ears; Site_1: Xba I; Site_2: Xba I;
The vector was digested with XbaI and one nucleotide was
added by fill in the recessive 3' end. The genomic DNA
was nebulized, end repaired, adaptor ligated and size
fractionated using sephadex. The resulting fragments were
between 0.8 and 3 kb and were cloned into the vector
(.x/y reads in M13mp19, b/g reads in pUC19). The same
ligation was transformed in either JM107 or DH5a."

BASE COUNT      161 a      71 c      57 g      119 t
ORIGIN
Query Match          64.0%; Score 16; DB 17; Length 408;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCCAGGACCAAGAA 16
|||||
Db 384 GTGCCAGGACCAAGAA 399

RESULT 18
BH876286          447 bp      DNA      linear      GSS 05-AUG-2002
LOCUS             hr30b07.bl WGS-ZmaysF (JM107 adapted methyl filtered) Zea mays
DEFINITION        genomic clone hr30b07 5', DNA sequence.

```

```

ACCESSION         BH876286
VERSION           BH876286.1 GI:22112183
KEYWORDS          GSS.
SOURCE            Zea mays.
ORGANISM          Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 447)
Rabinowicz,P.D., O'Shaughnessy,A.L., Balijs,V., Dedhia,N.,
Katzenburger,F., King,L., Miller,B., Muller,S., Nascimento,L.,
Zutavern,T., McCombie,W.R. and Martienssen,R.A.
Genomic shotgun sequences from Zea mays (methyl-filtered)
Unpublished (2002)
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: hr30 row: b column: 07
Seq primer: -21M13UnivFwd
Class: Shotgun
High quality sequence stop: 447.
Location/Qualifiers
1. 447
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="hr30b07"
/clone_host="JM107 or DH5a"
/label="Organ: immature ears; Site_1: Xba I; Site_2: Xba I;
The vector was digested with XbaI and one nucleotide was
added by fill in the recessive 3' end. The genomic DNA
was nebulized, end repaired, adaptor ligated and size
fractionated using sephadex. The resulting fragments were
between 0.8 and 3 kb and were cloned into the vector
(.x/y reads in M13mp19, b/g reads in pUC19). The same
ligation was transformed in either JM107 or DH5a."

BASE COUNT      155 a      86 c      65 g      137 t
ORIGIN
Query Match          64.0%; Score 16; DB 17; Length 447;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGACCAAGAA 16
|||||
Db 350 TGTGCCAGGACCAAGAA 365

RESULT 19
AA126284          449 bp      mRNA      linear      EST 23-DEC-1997
LOCUS             zm79e02.s1 Stratagene neuroepithelium (#937231) Homo sapiens cDNA
DEFINITION        clone IMAGE:531866 3', mRNA sequence.
ACCESSION         AA126284
VERSION           AA126284.1 GI:1685940
KEYWORDS          EST.
SOURCE            human.
ORGANISM          Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 449)
Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
Chisoe,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W., Hawkins
,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
,Schellenberg,K., Soares,M.B., Tran,F., Thierry-Mieg,J., Trevaskis,E.,
Underwood,K., Wohlmann,P., Waterston,R., Wilson,K. and Marra,M.
Generation and analysis of 280,000 human expressed sequence tags

```


JOURNAL
MEDLINE
COMMENT

Genome Res. 6 (9), 807-828 (1996)
97044478
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
WARNING: There is evidence that suggests that the 384-well parent plate of this clone contains both human and mouse derived clones. Thus, the origin of this clone is uncertain. This caution should be kept in mind should you use this clone.

This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 1893 Std Error: 0.00
Seq primer: -40M13 fwd. from Amersham
High quality sequence stop: 353.

FEATURES
source

```
1. .449
/organism="Homo sapiens"
/db_xref="GDB:3921378"
/db_xref="taxon:9606"
/clone="IMAGE:531866"
/dev_stage="Ntera-2/RA neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. NT2 cells (Ntera-2/ci.D1) induced with Retinoic Acid for 24 hours. Average insert size: 1.5 Kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTT 3'"
BASE COUNT      128 a      90 c      120 g      110 t      1 others
ORIGIN
```

```
Query Match      64.0%; Score 16; DB 9; Length 449;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      7 AGGAACCAAGAAATTGA 22
|||||
Db      418 AGGAACCAAGAAATTGA 433
```

```
RESULT 20
AI325084/c
LOCUS
DEFINITION      464 bp mRNA linear EST 23-DEC-1998
mq95f12.x1 Stratagene mouse heart (#937316) Mus musculus cDNA clone
IMAGE:586511 3' similar to TR:P70441 P70441 PROTEIN CO-FACTOR. ;
```

```
ACCESSION      AI325084
VERSION
KEYWORDS
SOURCE
house mouse.
ORGANISM
Mus musculus
```

```
REFERENCE
1 (bases 1 to 464)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
```

```
TITLE
The WashU-BHMI Mouse EST Project
JOURNAL
COMMENT
Contact: Marra M/Mouse EST Project
WashU-BHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
```

This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:361159

This clone was previously sequenced on the 5' end only, this new data is from the 3' end

Possible reversed clone: similarity on wrong strand
High quality sequence stop: 454.

FEATURES
source

```
Location/Qualifiers
1. .464
/organism="Mus musculus"
/strain="NIH Swiss"
/db_xref="taxon:10090"
/clone="IMAGE:586511"
/clone_lib="Stratagene mouse heart (#937316)"
/sex="pooled"
/tissue_type="heart"
/dev_stage="13 day embryos"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: heart; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. 93 pooled NIH/Swiss 13 day embryo hearts. Average insert size: 1.0 Kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTT 3'"
BASE COUNT      104 a      141 c      114 g      105 t
ORIGIN
```

```
Query Match      64.0%; Score 16; DB 9; Length 464;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 TGTGCGAGCAACACAGA 16
|||||
Db      110 TGTGCGAGCAACACAGA 95
```

```
RESULT 21
AI747269/c
LOCUS
DEFINITION      468 bp mRNA linear EST 22-JUN-1999
u114c03.x1 Sugano mouse embryo mewa Mus musculus cDNA clone
IMAGE:2076772 3', mRNA sequence.
```

```
ACCESSION      AI747269
VERSION
KEYWORDS
SOURCE
EST.
house mouse.
ORGANISM
Mus musculus
```

```
REFERENCE
1 (bases 1 to 468)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
```

```
TITLE
The WashU-NCI Mouse EST Project 1999
JOURNAL
COMMENT
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
```

This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:995208

Seq primer: custom primer used
High quality sequence stop: 437.

FEATURES
source

```
Location/Qualifiers
1. .468
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:2076772"
```

```

/clone_lib="Sugano mouse embryo mewa"
/dev_stage="embryo, 14 dpc"
/lab_host="DH10B"
/notes="Vector: pME18S-FL3; Site_1: DraIII (CACTGTGTC);
Site_2: DraIII (CACCATGTC); 1st strand cDNA was primed
with an oligo(dT) primer [ATGGCCCTTTTTTTTTTTTTT];
double-stranded cDNA was ligated to a DraIII adaptor
[TTTGGCTACTGG], digested and cloned into distinct DraIII
sites of the pME18S-FL3 vector (5' site CACTGTGTC, 3' site
CACCATGTC). XhoI should be used to isolate the cDNA
insert. Size selection was performed to exclude fragments
<1.5kb. Library constructed by Dr. Sumio Sugano
(University of Tokyo Institute of Medical Science).
Custom primers for sequencing: 5' end primer
CTTCTGCTAAAGCTGCG and 3' end primer
CGACCTGCAGCTCGACACA."
BASE COUNT      151 a   82 c   70 g   164 t   1 others
ORIGIN

Query Match      64.0%; Score 16; DB 9; Length 468;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 AGGACCAAGAAATTTA 22
|||||
Db 249 AGGACCAAGAAATTTA 234

RESULT 22
BH879621/c
LOCUS      BH879621
DEFINITION ht47f08-g1 WGS-ZmaysF (JM107 adapted methyl filtered) Zea mays
ACCESSION  BH879621
VERSION     BH879621.1 GI:22115518
KEYWORDS    GSS.
SOURCE      Zea mays.
ORGANISM    Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS    Rabinowicz P.D., O'Shaughnessy A.L., Balijs V., Dedhia N.,
Katzenburger F., King L., Miller B., Muller S., Nascimento L.,
Zucaveri T., McCombie W.R. and Martienssen R.A.
Genomic shotgun sequences from Zea mays (methyl-filtered)
Unpublished (2002)
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: ht47 row: f column: 08
Seq primer: -2L13UnivFwd
Class: Shotgun
High quality sequence stop: 471.
FEATURES
source
1. .471
Location/Qualifiers
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ht47f08"
/clone_lib="WGS-ZmaysF (JM107 adapted methyl filtered)"
/lab_host="JM107 or DH5a"
/notes="Organ: immature ears; Site_1: Xba I; Site_2: Xba I;
The vector was digested with XbaI and one nucleotide was
added by fill in in the recessive 3' end. The genomic DNA
was nebulized, end repaired, adaptor ligated and size
fractionated using sephadex. The resulting fragments were
between 0.8 and 3 kb and were cloned into the vector
(.x/y reads in M13mp19, .b/g reads in pUC19). The same
ligation was transformed in either JM107 or DH5a."
BASE COUNT      173 a   72 c   68 g   156 t
ORIGIN

Query Match      64.0%; Score 16; DB 17; Length 471;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTGCCAGGAACACAGA 16
|||||
Db 288 TGTGCCAGGAACACAGA 273

RESULT 23
B98791/c
LOCUS      B98791
DEFINITION CIT-HSP-2280N3.TR CIT-HSP Homo sapiens genomic clone 2280N3, DNA
sequence.
ACCESSION  B98791
VERSION     B98791.1 GI:3026015
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS    Adams M.D., Rounsley S.D., Zhao S., Field C.E., Bass S., Linher K.,
Golden K., Berry K., Granger D., Suh E., Wible C., Shizuya H.,
Simon M. and Venter J.C.
Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
Unpublished (1998)
Other_GSSs: CIT-HSP-2280N3.TF
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.
FEATURES
source
1. .489
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="GDB:7146965"
/db_xref="taxon:9606"
/clone="2280N3"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/notes="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT      130 a   97 c   106 g   156 t
ORIGIN

Query Match      64.0%; Score 16; DB 17; Length 489;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTGCCAGGAACACAGA 16
|||||
Db 462 TGTGCCAGGAACACAGA 447

RESULT 24
BF610439/c
LOCUS      BF610439
DEFINITION NXSI_058_E06_F NXSI (Nsf Xylem Side wood Inclined) Pinus taeda CDNA
clone NXSI_058_E06_5', mRNA sequence.
505 bp mRNA linear EST 14-DEC-2000

```

```

ACCESSION      BF610439
VERSION        BF610439.1 GI:11779094
KEYWORDS       EST
SOURCE         Lobolly pine.
ORGANISM       Pinus taeda

REFERENCE      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
AUTHORS        1 (bases 1 to 505)
TITLE          Sederoff, R.
JOURNAL        Molecular Basis of Wood Formation in the Pine Megagenome
COMMENT        Unpublished (2000)
               Contact: Johnson, Arthur
               North Carolina State University
               Tel: 919 515 7800
               Fax: 919 515 7801
               Email: ajohnson@unity.ncsu.edu
               Seq primer: T3.

FEATURES       source
               Location/Qualifiers
                 1..505
                   /organism="Pinus taeda"
                   /strain="Coastal plain lobolly pine from North Carolina"
                   /db_xref="taxon:3352"
                   /clone_lib="NXSI_058_E06"
                   /clone="NXSI_058_E06"
                   /tissue_type="Xylem"
                   /cell_type="Side"
                   /dev_stage="Juvenile"
                   /lab_host="XLI-Blue"
                   /note="Vector: Bluescript SK; Site.1: Eco RI; Site.2: XhoI
                   : The library is from early (spring) wood, taken from
                   : three six-year old trees (three different genotypes), in
                   : the juvenile phase. These trees were induced to form side
                   : wood by bending to a 45 degree angle and tying them to the
                   : ground. Differentiating xylem was harvested from the sides
                   : of the inclined stems, and a mixture of all three
                   : genotypes was used for the library. oligo-dt primed cDNA
                   : was directionally cloned into the EcoRI-XhoI Bluescript SK
                   : vector arms. NOTE: The sequences contain a 'cDNA adapter'
                   : between the EcoRI site and the start of the EST. The
                   : adapter sequence is 'AATCGGACGAG'."
                 134 a 114 c 111 g 130 t 16 others
BASE COUNT
ORIGIN

Query Match      64.0%; Score 16; DB 12; Length 505;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GCCAGGAACAGAAAT 19
    |||||
Db 181 GCCAGGAACAGAAAT 166

RESULT 25
BM725916
LOCUS          BM725916
DEFINITION    UI-E-EJ0-aig-i-22-0-UI.r1 UI-E-EJ0 Homo sapiens cDNA clone
               533 bp mRNA linear EST 01-MAR-2002
               UI-E-EJ0-aig-i-22-0-UI 5', mRNA sequence.
ACCESSION     BM725916
VERSION       BM725916
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 533)
AUTHORS       Bernaldo, M.F., Lennon, G. and Soares, M.B.
TITLE         Normalization and subtraction: two approaches to facilitate gene
               discovery
JOURNAL        Genome Res. 6 (9), 791-806 (1996)
MEDLINE        97044477
COMMENT       Contact: Soares, MB
               Program for Rat Gene Discovery and Mapping
               University of Iowa

ACCESSION      BF610439
VERSION        BF610439.1 GI:11779094
KEYWORDS       EST
SOURCE         Lobolly pine.
ORGANISM       Pinus taeda

REFERENCE      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
AUTHORS        1 (bases 1 to 505)
TITLE          Sederoff, R.
JOURNAL        Molecular Basis of Wood Formation in the Pine Megagenome
COMMENT        Unpublished (2000)
               Contact: Johnson, Arthur
               North Carolina State University
               Tel: 919 515 7800
               Fax: 919 515 7801
               Email: ajohnson@unity.ncsu.edu
               Seq primer: T3.

FEATURES       source
               Location/Qualifiers
                 1..505
                   /organism="Pinus taeda"
                   /strain="Coastal plain lobolly pine from North Carolina"
                   /db_xref="taxon:3352"
                   /clone_lib="NXSI_058_E06"
                   /clone="NXSI_058_E06"
                   /tissue_type="Xylem"
                   /cell_type="Side"
                   /dev_stage="Juvenile"
                   /lab_host="XLI-Blue"
                   /note="Vector: Bluescript SK; Site.1: Eco RI; Site.2: XhoI
                   : The library is from early (spring) wood, taken from
                   : three six-year old trees (three different genotypes), in
                   : the juvenile phase. These trees were induced to form side
                   : wood by bending to a 45 degree angle and tying them to the
                   : ground. Differentiating xylem was harvested from the sides
                   : of the inclined stems, and a mixture of all three
                   : genotypes was used for the library. oligo-dt primed cDNA
                   : was directionally cloned into the EcoRI-XhoI Bluescript SK
                   : vector arms. NOTE: The sequences contain a 'cDNA adapter'
                   : between the EcoRI site and the start of the EST. The
                   : adapter sequence is 'AATCGGACGAG'."
                 134 a 114 c 111 g 130 t 16 others
BASE COUNT
ORIGIN

Query Match      64.0%; Score 16; DB 12; Length 505;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GCCAGGAACAGAAAT 19
    |||||
Db 181 GCCAGGAACAGAAAT 166

RESULT 25
BM725916
LOCUS          BM725916
DEFINITION    UI-E-EJ0-aig-i-22-0-UI.r1 UI-E-EJ0 Homo sapiens cDNA clone
               533 bp mRNA linear EST 01-MAR-2002
               UI-E-EJ0-aig-i-22-0-UI 5', mRNA sequence.
ACCESSION     BM725916
VERSION       BM725916
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 533)
AUTHORS       Bernaldo, M.F., Lennon, G. and Soares, M.B.
TITLE         Normalization and subtraction: two approaches to facilitate gene
               discovery
JOURNAL        Genome Res. 6 (9), 791-806 (1996)
MEDLINE        97044477
COMMENT       Contact: Soares, MB
               Program for Rat Gene Discovery and Mapping
               University of Iowa

```

451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: mssoares@blue.weeg.uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse.

```

FEATURES       source
               Location/Qualifiers
                 1..533
                   /organism="Homo sapiens"
                   /db_xref="taxon:9606"
                   /clone="UI-E-EJ0-aig-i-22-0-UI"
                   /clone_lib="UI-E-EJ0"
                   /tissue_type="fetal eyes, lens, eye anterior segment,
                   optic nerve, retina, Retina Foveal and Macular, RPE and
                   Choroid"
                   /dev_stage="fetal and adult"
                   /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                   /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                   modified polylinker; Site.1: EcoR I; Site.2: Not I;
                   UI-E-EJ0 is a subcloned cDNA library constructed
                   according to Bernaldo, Lennon and Soares, Genome Research,
                   6:791-806, 1996. First strand cDNA synthesis was primed,
                   with an oligo-dT primer containing a Not I site. Double
                   stranded cDNA was ligated to an EcoR I adaptor, digested
                   with Not I, and cloned directionally into pT7T3-Pac
                   vector. The oligonucleotide used to prime the synthesis of
                   first-strand cDNA contains a library tag sequence that is
                   located between the Not I site and the (dT)18 tail. The
                   sequence tags for this library are: fetal eyes, AGAATCAGAA
                   ; lens, CGATAGGCGA; eye anterior segment, AATGCCGCGAT;
                   optic nerve, CCATTAGTGT; retina, CCGCG; Retina Foveal and
                   Macular, GTCC; RPE and Choroid, ACCTA. This library was
                   created for the program, Gene Discovery in the Visual
                   System, supported by National Eye Institute (NEI)."
                 96 a 175 c 121 g 141 t
BASE COUNT
ORIGIN

```

```

Query Match      64.0%; Score 16; DB 14; Length 533;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTGCCAGGACGACGAA 17
    |||||
Db 105 GTGCCAGGACGACGAA 120

```

```

RESULT 26
BQ701598/c
LOCUS          BQ701598
DEFINITION    NXSI_111_E01_F NXSI (Nsf Xylem Side wood Inclined) Pinus taeda cDNA
               clone NXSI_111_E01 5', mRNA sequence.
ACCESSION     BQ701598
VERSION       BQ701598
KEYWORDS      EST.
SOURCE        lobolly pine.
ORGANISM      Pinus taeda
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE     1 (bases 1 to 543)
AUTHORS       Sederoff, R.
TITLE         Molecular Basis of Wood Formation in the Pine Megagenome
JOURNAL        Unpublished (2000)
COMMENT       Contact: Johnson, Arthur
               North Carolina State University
               Tel: 919 515 7800
               Fax: 919 515 7801
               Email: ajohnson@unity.ncsu.edu
               Seq primer: T3.

```

FEATURES
source

Location/Qualifiers
1. 543
/organism="Pinus taeda"
/db_xref="taxon:3352"
/clone="NXSI_111.E01"
/clone_lib="NXSI (Nsf Xylem Side wood Inclined)"
/tissue_type="Xylem"
/cell_type="Side"
/dev_stage="Juvenile"
/lab_host="XLI-Blue"
/note="Vector: Bluescript SK; Site_1: Eco RI; Site_2: XhoI
; The library is from early (spring) wood, taken from
three six-year old trees (three different genotypes), in
the juvenile phase. These trees were induced to form side
wood by bending to a 45 degree angle and tying them to the
ground. Differentiating xylem was harvested from the sides
of the inclined stems, and a mixture of all three
genotypes was used for the library. oligo-dt primed cDNA
was directionally cloned into the EcoRI-XhoI Bluescript SK
vector arms. NOTE: The sequences contain a 'cDNA adapter'
between the EcoRI site and the start of the EST. The
adapter sequence is 'AATTCGCACGAG'."
BASE COUNT 148 a 127 c 118 g 139 t 11 others
ORIGIN

Query Match 64.0%; Score 16; DB 14; Length 543;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 GCCAGGAACCAAGAT 19
Db 241 GCCAGGAACCAAGAT 226
|||||

RESULT 27
BH876388
LOCUS 551 bp DNA linear GSS 05-AUG-2002
DEFINITION hr30906.g1 WGS-ZmaysF (JM107 adapted methyl filtered) Zea mays
ACCESSION BH876388
VERSION BH876388.1 GI:22112285
KEYWORDS GSS.
SOURCE Zea mays.
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
Clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 551)
Rabinowicz, P.D., O'Shaughnessy, A.L., Ballja, V., Dedhia, N.,
Katzenburger, F., King, L., Miller, B., Muller, S., Nascimento, L.,
Zutavern, T., McCombie, W.R. and Martienssen, R.A.
Genomic shotgun sequences from Zea mays (methyl-filtered)
Unpublished (2002)
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: hr30 row: g column: 06
Seq primer: -21M13UnivFwd
Class: shotgun
High quality sequence stop: 551.

FEATURES
source

Location/Qualifiers
1. 551
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="hr30906"
/clone_lib="WGS-ZmaysF (JM107 adapted methyl filtered)"
/lab_host="JM107 or DH5a"

/note="Organ: immature ears; Site_1: Xba I; Site_2: Xba I;
The vector was digested with XbaI and one nucleotide was
added by fill in in the recessive 3' end. The genomic DNA
was nebulized, end repaired, adaptor ligated and size
fractionated using sephadex. The resulting fragments were
between 0.8 and 3 kb and were cloned into the vector
(.x/y reads in M13mp19, .b/g reads in pUC19). The same
ligation was transformed in either JM107 or DH5a."
BASE COUNT 204 a 100 c 78 g 169 t
ORIGIN

Query Match 64.0%; Score 16; DB 17; Length 551;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TGTCGAGGAACCAAGA 16
Db 388 TGTCGAGGAACCAAGA 403
|||||

RESULT 28
BG039852/c
LOCUS 557 bp mRNA linear EST 24-JAN-2001
DEFINITION NXSI_104_G04_F NXSI (Nsf Xylem Side wood Inclined) Pinus taeda cDNA
clone NXSI_104_G04_5, mRNA sequence.
ACCESSION BG039852
VERSION BG039852.1 GI:12482533
KEYWORDS EST.
SOURCE loblolly pine.
ORGANISM Pinus taeda

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
1 (bases 1 to 557)
Sederoff, R.
Molecular Basis of Wood Formation in the Pine Megagenome
Unpublished (2000)
Contact: Johnson, Arthur
North Carolina State University
Tel: 919 515 7800
Fax: 919 515 7801
Email: ajohnson@unity.ncsu.edu
Seq primer: T3.

FEATURES
source

Location/Qualifiers
1. 557
/organism="Pinus taeda"
/strain="Coastal plain loblolly pine from North Carolina"
/db_xref="taxon:3352"
/clone="NXSI_104_G04"
/clone_lib="NXSI (Nsf Xylem Side wood Inclined)"
/tissue_type="Xylem"
/cell_type="Side"
/dev_stage="Juvenile"
/lab_host="XLI-Blue"
/note="Vector: Bluescript SK; Site_1: Eco RI; Site_2: XhoI
; The library is from early (spring) wood, taken from
three six-year old trees (three different genotypes), in
the juvenile phase. These trees were induced to form side
wood by bending to a 45 degree angle and tying them to the
ground. Differentiating xylem was harvested from the sides
of the inclined stems, and a mixture of all three
genotypes was used for the library. oligo-dt primed cDNA
was directionally cloned into the EcoRI-XhoI Bluescript SK
vector arms. NOTE: The sequences contain a 'cDNA adapter'
between the EcoRI site and the start of the EST. The
adapter sequence is 'AATTCGCACGAG'."
BASE COUNT 149 a 123 c 121 g 147 t 17 others
ORIGIN

Query Match 64.0%; Score 16; DB 12; Length 557;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 GCCAGGAACCAAGAT 19

```

|||||
Db 195 GCCAGGAACAGAAAT 180

RESULT 29
AW644504/c
LOCUS
DEFINITION
cm41h08.w1 Blackshear/Soares normalized Xenopus egg library xenopus
laevis cDNA clone PBX0140H08 5', mRNA sequence.
ACCESSION
AW644504
VERSION
AW644504.1 GI:7401893
KEYWORDS
EST.
SOURCE
African clawed frog.
ORGANISM
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus.
REFERENCE
1 (bases 1 to 574)
Blackshear, P.J., Lai, W.S., Thorn, J.M., Kennington, E.A., Staffa, N.G.,
Jr., Moore, D.T., Bouffard, G.G., Beckstrom-Sternberg, S.M., Touchman,
J.W., Bonaldo, M.F. and Soares, M.B.
The NIHES Xenopus maternal EST project: interim analysis of the
first 13,879 ESTs from unfertilized eggs
Gene 267 (1), 71-87 (2001)
21211403
Contact: Perry J. Blackshear
Office of Clinical Research and Laboratory of Signal Transduction
National Institute of Environmental Health Sciences
A2-05 NIHES, 101 Alexander Drive, Research Triangle Park, NC 27709,
USA
Tel: 919 541-4899
Fax: 919 541-4571
Email: black009@niehs.nih.gov
Clone is available through Research Genetics, Inc., 2130 Memorial
Parkway, Huntsville, AL 35901
phone 800-533-4363 ext.cdn, fax 256-536-9016 att:cdna, email
cdna@resgen.com
DNA Sequencing and analyses performed by National Institutes of
Health Intramural Sequencing Center (NISC).
PCR Primers
FORWARD: TGTAAACGCGGCGCAGT
BACKWARD: CAGGAACAGCTATGACC
Plate: 0140 row: H column: 08
Seq primer: T7 primer.
FEATURES
Source
Location/Qualifiers
1..574
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="PBX0140H08"
/clone_lib="Blackshear/Soares normalized Xenopus egg
library"
/sex="female"
/tissue_type="unfertilized egg"
/cell_type="unfertilized egg"
/dev_stage="unfertilized egg"
/lab_host="DH10B"
/note="Vector: pT73-Pac; Site1: EcoRI; Site2: NotI;
PolyA-selected mRNA was prepared from unfertilized Xenopus
laevis eggs. The library was constructed in the vector
pT73-Pac as described in Bonaldo, M.F., Lennon, G. and
Soares, M.B. 'Normalization and subtraction: two
approaches to facilitate gene discovery', Genome Research
6:791-806, 1996. The first strand synthesis used a
NotI-dt18 primer. Double stranded cDNAs were ligated to
EcoRI adapters, digested with NotI, and directionally
cloned into the NotI and EcoRI-digested pT73-Pac vector.
The library contained approximately 7.2 x 10^5
recombinants, with average insert sizes of 1-1.5 kb."
BASE COUNT
165 a 109 c 116 g 184 t
ORIGIN

Query Match
Best Local Similarity 64.0%; Score 16; DB 10; Length 574;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 31
BH876535
LOCUS
DEFINITION
hr31g03.b1 WGS-ZmaysF (JM107 adapted methyl filtered) Zea mays
genomic clone hr31g03 5', DNA sequence.
ACCESSION
BH876535
VERSION
BH876535.1 GI:22112432
KEYWORDS
GSS.
ORGANISM
Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 581)
Rabinowicz, P.D., O'Shaughnessy, A.L., Ballja, V., Dedhia, N.,
Katzenburger, F., King, L., Miller, B., Muller, S., Nascimento, L.,
Zutavern, T., McCombie, W.R. and Martienssen, R.A.
Genomic shotgun sequences from Zea mays (methyl-filtered)
Unpublished (2002)
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: hr30 row: g column: 06
Seq primer: -21M13UnivFwd
Class: shotgun
High quality sequence stop: 581.
FEATURES
Source
Location/Qualifiers
1..581
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="hr30g06"
/clone_lib="WGS-ZmaysF (JM107 adapted methyl filtered)"
/lab_host="JM107 or DH5a"
/note="Organ: immature ears; Site1: Xba I; Site2: Xba I;
The vector was digested with XbaI and one nucleotide was
added by fill in in the recessive 3' end. The genomic DNA
was nebulized, end repaired, adaptor ligated and size
fractionated using sephadex. The resulting fragments were
between 0.8 and 3 kb and were cloned into the vector
(.x/y reads in M13mp19, .b/g reads in pUC19). The same
ligation was transformed in either JM107 or DH5a."
BASE COUNT
178 a 83 c 97 g 223 t
ORIGIN

Query Match
Best Local Similarity 64.0%; Score 16; DB 17; Length 581;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 31
BH876535
LOCUS
DEFINITION
hr31g03.b1 WGS-ZmaysF (JM107 adapted methyl filtered) Zea mays
genomic clone hr31g03 5', DNA sequence.
ACCESSION
BH876535
VERSION
BH876535.1 GI:22112432
KEYWORDS
GSS.
ORGANISM
Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 581)
Rabinowicz, P.D., O'Shaughnessy, A.L., Ballja, V., Dedhia, N.,
Katzenburger, F., King, L., Miller, B., Muller, S., Nascimento, L.,
Zutavern, T., McCombie, W.R. and Martienssen, R.A.
Genomic shotgun sequences from Zea mays (methyl-filtered)
Unpublished (2002)
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: hr30 row: g column: 06
Seq primer: -21M13UnivFwd
Class: shotgun
High quality sequence stop: 581.
FEATURES
Source
Location/Qualifiers
1..581
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="hr30g06"
/clone_lib="WGS-ZmaysF (JM107 adapted methyl filtered)"
/lab_host="JM107 or DH5a"
/note="Organ: immature ears; Site1: Xba I; Site2: Xba I;
The vector was digested with XbaI and one nucleotide was
added by fill in in the recessive 3' end. The genomic DNA
was nebulized, end repaired, adaptor ligated and size
fractionated using sephadex. The resulting fragments were
between 0.8 and 3 kb and were cloned into the vector
(.x/y reads in M13mp19, .b/g reads in pUC19). The same
ligation was transformed in either JM107 or DH5a."
BASE COUNT
178 a 83 c 97 g 223 t
ORIGIN

```

SOURCE Zea mays.
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 596)
AUTHORS Rabinowicz, P.D., O'Shaughnessy, A.L., Ballja, V., Dedhia, N., King, L., Miller, B., Muller, S., Nascimben, L., Katzenburger, F., McCombie, W.R. and Martienssen, R.A.

TITLE Genomic shotgun sequences from Zea mays (methyl-filtered)
JOURNAL Unpublished (2002)
COMMENT Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: hr31 row: g column: 03
Seq primer: -21M13univFwd
Class: shotgun
High quality sequence stop: 596.

FEATURES source
1. 596
Location/Qualifiers
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="hr31g03"
/lab_host="JM107 or DH5a"
/note="Organ: immature ears; Site_1: Xba I; Site_2: Xba I; The vector was digested with XbaI and one nucleotide was added by fill in in the recessive 3' end. The genomic DNA was nebulized, end repaired, adaptor ligated and size fractionated using sephadex. The resulting fragments were between 0.8 and 3 kb and were cloned into the vector (-x/y reads in M13mp19, b/y reads in pUC19). The same ligation was transformed in either JM107 or DH5a."

BASE COUNT 165 a 109 c 105 g 217 t
ORIGIN

Query Match 64.0%; Score 16; DB 17; Length 596;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCACGACGACGACA 16
Db 481 TGTGCACGACGACGACA 496

RESULT 32
AW043028/c
LOCUS
DEFINITION AW043028 Pine Triplex shoot tip library Pinus taeda cDNA clone
ACCESSION AW043028
VERSION AW043028.1 GI:5903557
KEYWORDS EST.
SOURCE loblolly pine.
ORGANISM Pinus taeda

REFERENCE 1 (bases 1 to 601)
AUTHORS Whetten, R.W., Kinlaw, C.S., Retzel, E. and Sederoff, R.R.
TITLE The Pine Gene Discovery Project
JOURNAL Unpublished (1999)
COMMENT Contact: Ross Whetten
Forest Biotechnology Group
North Carolina State University
Dept. of Forestry, NC State University, 6113 Jordan Hall, Raleigh, NC, 27695-8008
Tel: 919-515-7800
Fax: 919-515-7801

Query Match 64.0%; Score 16; DB 17; Length 596;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCACGACGACGACA 16
Db 481 TGTGCACGACGACGACA 496

RESULT 32
AW043028/c
LOCUS
DEFINITION AW043028 Pine Triplex shoot tip library Pinus taeda cDNA clone
ACCESSION AW043028
VERSION AW043028.1 GI:5903557
KEYWORDS EST.
SOURCE loblolly pine.
ORGANISM Pinus taeda

REFERENCE 1 (bases 1 to 601)
AUTHORS Whetten, R.W., Kinlaw, C.S., Retzel, E. and Sederoff, R.R.
TITLE The Pine Gene Discovery Project
JOURNAL Unpublished (1999)
COMMENT Contact: Ross Whetten
Forest Biotechnology Group
North Carolina State University
Dept. of Forestry, NC State University, 6113 Jordan Hall, Raleigh, NC, 27695-8008
Tel: 919-515-7800
Fax: 919-515-7801

Email: rosswhet@unity.ncsu.edu
Seq primer: 5' lambda Triplex2 Sequencing Primer.
Location/Qualifiers
1. 601
/organism="Pinus taeda"
/db_xref="taxon:3352"
/clone="ST28D06"
/clone_lib="Pine Triplex shoot tip library"
/lab_host="E. coli BM25.8"
/note="Organ: shoot tips; Vector: Lambda Triplex; Site_1: SfiI (A); Site_2: SfiI (B); Shoot tips (approx. 2 cm from apex) were collected during the spring, frozen and used for mRNA isolation. The SMART-PCR method (Clontech) was used to prepare a library from 1 ug total RNA, using the Lambda Triplex vector. Plasmid subclones in pTriplex were recovered by cre-lox excision in E. coli strain BM25.8 and sequenced from the 5' end."

BASE COUNT 166 a 133 c 118 g 170 t 14 others
ORIGIN

Query Match 64.0%; Score 16; DB 10; Length 601;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GCCAGGAACGAGAAAT 19
Db 512 GCCAGGAACGAGAAAT 497

RESULT 33
AW043050/c
LOCUS
DEFINITION AW043050 Pine Triplex shoot tip library Pinus taeda cDNA clone
ACCESSION ST28F07, mRNA sequence.
VERSION ST28F07, mRNA sequence.
KEYWORDS EST.
SOURCE loblolly pine.
ORGANISM Pinus taeda

REFERENCE 1 (bases 1 to 601)
AUTHORS Whetten, R.W., Kinlaw, C.S., Retzel, E. and Sederoff, R.R.
TITLE The Pine Gene Discovery Project
JOURNAL Unpublished (1999)
COMMENT Contact: Ross Whetten
Forest Biotechnology Group
North Carolina State University
Dept. of Forestry, NC State University, 6113 Jordan Hall, Raleigh, NC, 27695-8008
Tel: 919-515-7800
Fax: 919-515-7801
Email: rosswhet@unity.ncsu.edu
Seq primer: 5' lambda Triplex2 Sequencing Primer.
Location/Qualifiers
1. 601
/organism="Pinus taeda"
/db_xref="taxon:3352"
/clone="ST28F07"
/clone_lib="Pine Triplex shoot tip library"
/note="Organ: shoot tips; Vector: Lambda Triplex; Site_1: SfiI (A); Site_2: SfiI (B); Shoot tips (approx. 2 cm from apex) were collected during the spring, frozen and used for mRNA isolation. The SMART-PCR method (Clontech) was used to prepare a library from 1 ug total RNA, using the Lambda Triplex vector. Plasmid subclones in pTriplex were recovered by cre-lox excision in E. coli strain BM25.8 and sequenced from the 5' end."

BASE COUNT 176 a 135 c 121 g 169 t
ORIGIN

Query Match 64.0%; Score 16; DB 10; Length 601;

Best Local Similarity 100.0%; Pred. No. 95;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GCCAGGACCAAGAAAT 19
|||||
Db 514 GCCAGGACCAAGAAAT 499

RESULT 34
AW043091

LOCUS ST29C02 Pine Triplex shoot tip library Pinus taeda cDNA clone
DEFINITION ST29C02, mRNA sequence.

ACCESSION AW043091
VERSION AW043091.1 GI:5903620
KEYWORDS EST.
SOURCE loblolly pine.
ORGANISM Pinus taeda

REFERENCE 1 Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinus; Pinus.
AUTHORS Whetten, R.W., Kinlaw, C.S., Retzel, B. and Sederoff, R.R.
TITLE The Pine Gene Discovery Project
JOURNAL Unpublished (1999)
COMMENT Contact: Ross Whetten
Forest Biotechnology Group
North Carolina State University
Dept. of Forestry, NC State University, 6113 Jordan Hall, Raleigh
, NC, 27695-8008
Tel: 919-515-7800
Fax: 919-515-7801
Email: rosswhetten@unity.ncsu.edu

Seq primer: 5' lambda Triplex2 sequencing Primer.
Location/Qualifiers

FEATURES
source

1..601
/organism="Pinus taeda"
/db_xref="taxon:3352"
/clone="ST29C02"
/clone_lib="Pine Triplex shoot tip library"
/lab_host="E. coli BM25.8"
/note="Organ: shoot tips; Vector: Lambda Triplex; Site_1:
SfiI (A); Site_2: SfiI (B); Shoot tips (approx. 2 cm from
apex) were collected during the spring, frozen and used
for mRNA isolation. The SMART-PCR method (Clontech) was
used to prepare a library from 1 ug total RNA, using the
Lambda Triplex vector. Plasmid subclones in pTriplex were
recovered by cre-lox excision in E. coli strain BM25.8 and
sequenced from the 5' end."
BASE COUNT 187 a 134 c 135 g 141 t
ORIGIN 4 others

Query Match 64.0%; Score 16; DB 10; Length 601;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 GAACCAAGAAATTACA 24
|||||
Db 34 GAACCAAGAAATTACA 49

RESULT 35
AQ009748/c

LOCUS AQ009748
DEFINITION CIT-HSP-2280C6.TRB CIT-HSP Homo sapiens genomic clone 2280C6, DNA
sequence.

ACCESSION AQ009748
VERSION AQ009748.1 GI:3129133
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 608)

AUTHORS

TITLE
JOURNAL
COMMENT

Adams, M.D., Rounsley, S.D., Zhao, S., Field, C.E., Bass, S., Linher, K.,
Golden, K., Berry, K., Grainger, D., Suh, E., Wible, C., Shizuya, H.,
Simon, M. and Venter, J.C.
Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
Unpublished (1998)
Other GSSs: CIT-HSP-2280C6.TFB CIT-HSP-2280C6.TF

Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tcdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
source

1..608
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="GDB:7146704"
/db_xref="taxon:9606"
/clone="2280C6"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:
HindIII"

BASE COUNT 191 a 107 c 132 g 176 t 2 others
ORIGIN

Query Match 64.0%; Score 16; DB 17; Length 608;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGACACAGA 16
|||||
Db 313 TGTGCCAGGACACAGA 298

RESULT 36
BE600467

LOCUS PIL_96.D12.q1.A002 Pathogen induced 1 (PIL) Sorghum bicolor cDNA,
DEFINITION mRNA sequence.
ACCESSION BE600467
VERSION BE600467.1 GI:9855526
KEYWORDS EST.
SOURCE Sorghum bicolor
ORGANISM Sorghum bicolor

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 613)
AUTHORS Cordonnier-Pratt, M.-M., Gingle, A., Dean, R., Sudman, M. and Pratt
, L.H.
An EST database from Sorghum: pathogen-induced plants
Unpublished (2000)
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu

Sequences have been trimmed to exclude polyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.

Seq primer: PolyTMix
High quality sequence start: 14
High quality sequence stop: 608

```

POLYA=No.
FEATURES
    source
        Location/Qualifiers
            1..613
                /organism="Sorghum bicolor"
                /db_xref="taxon:4558"
                /clone_lib="Pathogen induced 1 (PII)"
                /note="Organ: Anthracnose-infected leaves from
                two-week-old sorghum plants 48 hr after inoculation;
                Vector: pBluescript II from lambda Zap II; Site 1: xhoI;
                Site 2: EcoRI; Two-week-old sorghum plants (BX 623
                cultivar) were infected with pathogen (isolate PM421 of
                Colletotrichum graminicola, which is a sorghum isolate).
                RNA was prepared from infected leaves harvested from 45
                seedlings 48 hours after inoculation. Note: young
                seedlings (2 weeks old) exhibit juvenile resistant
                reaction, which is an incompatible interaction. As they
                grow older (4 weeks or older), plants resume susceptibility
                to anthracnose disease. The library was made from poly-A
                RNA in the cloning vector lambda Zap II. Clones to be
                sequenced were prepared by mass excision. WARNING: While
                most or all ESTs are expected to derive from the host
                plant, no effort was made to eliminate ESTs deriving from
                the pathogen."
            228 a 104 c 94 g 187 t
            BASE COUNT
            ORIGIN
                228 a 104 c 94 g 187 t
                Query Match 64.0%; Score 16; DB 10; Length 613;
                Best Local Similarity 100.0%; Pred. No. 95;
                Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
                QY 1 TGTGCCAGGACACAGA 16
                Db 377 TGTGCCAGGACACAGA 392
                |||||||||||||||
                RESULT 37
                BG003085/c
                LOCUS
                DEFINITION
                ACCESSION
                VERSION
                KEYWORDS
                SOURCE
                ORGANISM
                Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                1 (bases 1 to 628)
                Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
                Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.,
                Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
                Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
                M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
                Simpson, A.J.
                Shotgun sequencing of the human transcriptome with ORF expressed
                sequence tags
                Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
                20202663
                Contact: Simpson A.J.G.
                Laboratory of Cancer Genetics
                Ludwig Institute for Cancer Research
                Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
                Brazil
                Tel: +55-11-2704922
                Fax: +55-11-2707001
                Email: asimpson@ludwig.org.br
                This sequence was derived from the FAPESP/LICR Human Cancer Genome
                Project. This entry can be seen in the following URL
                (http://www.ludwig.org.br/scripts/gethtml2.pl?l1=MR3&t2=MR3-GN0187-
                171100-011-f05&t3=2000-11-17&t4=1)
                Seq primer: puc 18 forward
                High quality sequence stop: 617.
                Location/Qualifiers
                    1..628
                FEATURES
                    source

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0187"
/dev_stage="Adult"
/note="Organ: Placenta_normal; Vector: puc18; Site 1: SmaI
; Site 2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
Application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."
129 a 154 c 197 g 148 t
BASE COUNT
ORIGIN
    129 a 154 c 197 g 148 t
    Query Match 64.0%; Score 16; DB 12; Length 628;
    Best Local Similarity 100.0%; Pred. No. 95;
    Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    QY 1 TGTGCCAGGACACAGA 16
    Db 507 TGTGCCAGGACACAGA 492
    |||||||||||||||
    RESULT 38
    BF511713/c
    LOCUS
    DEFINITION
    ACCESSION
    VERSION
    KEYWORDS
    SOURCE
    ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
    1 (bases 1 to 636)
    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
    National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
    Unpublished (1997)
    Tumor Gene Index
    Contact: Robert Strausberg, Ph.D.
    Email: cgaps-r@mail.nih.gov
    The sequence contained an oligo-dT track that was present in the
    oligonucleotide that was used to prime the synthesis of first
    strand cDNA and therefore this may represent a bonafide poly A
    tail. cDNA library Preparation: M.B. Soares Lab Clone distribution:
    NCI-CGAP clone distribution information can be found through the
    I.M.A.G.E. Consortium/LLNL at:
    www.bio.llnl.gov/bbrp/image/image.html The following repetitive
    elements were found in this cDNA sequence: 228-285, >MIR
    Seq primer: M13 Forward
    POLYA=Yes.
    Location/Qualifiers
        1..636
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone_lib="NCI-CGAP_Sub8"
        /lab_host="DH10B (Life Technologies)"
        /note="Vector: pT73D-Pac (Pharmacia) with a modified
        polylinker; Site 1: Not I; Site 2: Eco RI; NCI-CGAP_Sub8
        is a subtracted library derived from NCI-CGAP_Sub5. The
        NCI-CGAP_Sub8 library had 2.5 million recombinants. A
        single-stranded DNA preparation of NCI-CGAP_Sub5 was used
        as a tracer in a subtractive hybridization with a driver
        comprising: a pool of clones from NCI-CGAP_Sub5 (IMAGE
        clone ids 2732833-2737415, 3068040-3069191; 25% of the
        driver population), a pool of clones from NCI-CGAP_Sub4
        (IMAGE clone ids 2723592-2729326; 25% of the driver
        population), NCI-CGAP_Sub6 (pool AIF-AJU, IMAGE ids
        2728969-2733190; 25% of the driver population), and
        NCI-CGAP_Sub7 (IMAGE ids 3069192-3072238, 3081864-3084550
        ; 25% of the driver population). Subtraction was

```


performed as previously described [Bonaldo, Lennon & Soares (1996)]: Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery. Genome Research 6, 791-806.

BASE COUNT 187 a 135 c 119 g 194 t 1 others
 ORIGIN
 Query Match 64.0%; Score 16; DB 12; Length 636;
 Best Local Similarity 100.0%; Pred. No. 96;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 GCCAGCAACCAAGAAAT 19
 |||||
 Db 163 GCAGGACCAAGAAAT 148

RESULT 39
 BM678024/c
 LOCUS 663 bp mRNA linear EST 27-FEB-2002
 DEFINITION UI-E-EJ0-aig-1-22-0-UI.s1 UI-E-EJ0 Homo sapiens cDNA clone
 UI-E-EJ0-aig-1-22-0-UI 3', mRNA sequence.
 ACCESSION BM678024
 VERSION BM678024.1 GI:18987920
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 663)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 9704477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Forward
 POLYA=Yes.

FEATURES
 source
 1..663
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="UI-E-EJ0-aig-1-22-0-UI"
 /clone_lib="UI-E-EJ0"
 /tissue_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina Foveal and Macular, RPE and Choroid"
 /dev_stage="fetal and adult"
 /lab_host="DH10B (Life Technologies) (TI phage resistant)"
 /note="Organ: eye; Vector: pT7n3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-EJ0 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7n3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGAATCAAGA

; lens, CCATTAGCGA; eye anterior segment, AATGCCGAT; optic nerve, CCATTAGTG; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).
 TAG_LIB=UI-E-EJ0
 TAG_ISSUE=human retina
 TAG_SEQ=CCGCG"

BASE COUNT 162 a 159 c 196 g 145 t 1 others
 ORIGIN
 Query Match 64.0%; Score 16; DB 14; Length 663;
 Best Local Similarity 100.0%; Pred. No. 97;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTGCCAGGACCAAGAA 17
 |||||
 Db 550 GTGCCAGGACCAAGAA 535

RESULT 40
 AG133215
 LOCUS 667 bp DNA linear GSS 04-NOV-2001
 DEFINITION Pan troglodytes DNA, clone: PTB-145N04.F, genomic survey sequence.
 ACCESSION AG133215
 VERSION AG133215.1 GI:16662893
 KEYWORDS GSS.
 SOURCE

ORGANISM Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male
 BAC Library clone:PTB-145N04.F.
 Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.
 REFERENCE 1
 AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 TITLE BAC end sequences of Library PTB
 JOURNAL Unpublished

REFERENCE 2 (bases 1 to 667)
 AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:chimpesgsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)
 COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.
 PRIMERS
 Sequencing: -21M13

LIBRARY
 Vector : pKS145
 R.Site 1 : SacI
 R.Site 2 : SacI.
 Location/Qualifiers
 1..667
 /organism="Pan troglodytes"
 /db_xref="taxon:9598"
 /clone="PTB-145N04.F"
 /sex="male"
 /cell_type="lymphoblast"
 /clone_lib="PTB Chimpanzee Male BAC Library"

BASE COUNT 176 a 191 c 135 g 165 t
 ORIGIN
 Query Match 64.0%; Score 16; DB 17; Length 667;
 Best Local Similarity 100.0%; Pred. No. 97;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGACCAAGAA 16
 |||||
 Db 307 TGTGCCAGGACCAAGAA 322

RESULT 41
 BH882704
 LOCUS
 DEFINITION 686 bp DNA linear GSS 05-AUG-2002
 genomic clone hw39b01 5', DNA sequence.
 ACCESION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 686)
 Rabinowicz, P.D., O'Shaughnessy, A.L., Balija, V., Dedhia, N.,
 Katzenburger, F., King, L., Miller, B., Muller, S., Nascento, L.,
 Zutavern, T., McCombie, W.R., and Martienssen, R.A.
 Genomic shotgun sequences from Zea mays (methyl-filtered)
 Unpublished (2002)
 CONTACT: W. Richard McCombie
 Lita Annenberg Hazen Genome Sequencing Center
 Cold Spring Harbor Laboratory
 PO Box 100, Cold Spring Harbor, NY 11724, USA
 Tel: 516 367 8884
 Fax: 516 367 8874
 Email: mcombie@cshl.org
 Plate: hw39 row: b column: 01
 Seq primer: -21M13UnivFwd
 Class: Shotgun
 High quality sequence stop: 686.
 Location/Qualifiers
 1..686
 /organism="Zea mays"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone="hw39b01"
 /clone_lib="WGS-ZmaysF (JM107 adapted methyl filtered)"
 /lab_host="JM107 or DH5a"
 /note="Organ: immature ears; Site_1: Xba I; Site_2: Xba I;
 The vector was digested with XbaI and one nucleotide was
 added by fill in in the recessive 3' end. The genomic DNA
 was nebulized, end repaired, adaptor ligated and size
 fractionated using sephadex. The resulting fragments were
 between 0.8 and 3 kb and were cloned into the vector
 (X/Y reads in M13mpl9, B/g reads in pUC19). The same
 ligation was transformed in either JM107 or DH5a."
 243 a 109 c 121 g 213 t
 BASE COUNT
 ORIGIN
 Query Match 64.0%; Score 16; DB 17; Length 686;
 Best Local Similarity 100.0%; Pred. No. 97;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGTGCCAGGACACCA 16
 |||||
 Db 111 TGTGCCAGGACACCA 126
 RESULT 42
 AW340752/C
 LOCUS
 DEFINITION 698 bp mRNA linear EST 31-JAN-2000
 hd01g09.x1 Soares_NFL.T.GBC.S1.Homo sapiens cDNA clone
 IMAGE:2908288 3' similar to WP:707A9.5 CEL7214 ; mRNA sequence.
 ACCESION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 698)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cyapbs@mail.nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -400P from Gibco
 High quality sequence stop: 462.
 Location/Qualifiers
 1..698
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2908288"
 /clone_lib="Soares_NFL.T.GBC.S1"
 /lab_host="DH10B"
 /note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
 a modified polylinker; Site_1: Not I; Site_2: Eco RI;
 Equal amounts of plasmid DNA from three normalized
 libraries (fetal lung NDHL19W, testis NHT, and B-cell
 NCI-CGAP-GC81) were mixed, and ss circles were made in
 vitro. Following HAP purification, this DNA was used as
 tracer in a subtractive hybridization reaction. The driver
 was PCR-amplified cDNAs from pools of 5,000 clones made
 from the same 3 libraries. The pools consisted of
 I.M.A.G.E. clones 297480-302087, 682632-687239,
 726408-728711, and 729096-731399. Subtraction by Bento
 Soares and M. Fatima Bonaldo."
 207 a 153 c 135 g 201 t
 BASE COUNT
 ORIGIN
 Query Match 64.0%; Score 16; DB 10; Length 698;
 Best Local Similarity 100.0%; Pred. No. 98;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 GCCAGGACCGAATAAT 19
 |||||
 Db 145 GCCAGGACCGAATAAT 130
 RESULT 43
 AQ745078
 LOCUS
 DEFINITION 752 bp DNA linear GSS 16-JUL-1999
 HS.5506.AL.E07.SP6.RPC1-11 Human Male BAC Library Homo sapiens
 genomic clone Plate=1082 Col=13 Row=L, DNA sequence.
 ACCESION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 752)
 Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
 Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
 Hood, L.
 Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 99380589
 CONTACT: Mahairas G.G., Wallace J.C., Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPC1-11. For BAC
 library availability, please contact Pieter de Jong
 (pieterdejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering.bac.htm>)
 or from Resear h Genetics (info@resgen.com). BAC end Web Server:

http://www.htsc.washington.edu
 Plate: 1082 row: i column: 13
 Seq primer: SP6
 Class: BAC ends
 High quality sequence stop: 752.
 Location/Qualifiers

FEATURES

source
 1..752
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="Plate:1082 Col=13 Row=I"
 /clone_lib="RPC1-11 Human Male BAC Library"
 /sex="male"
 /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
 Male blood DNA was isolated from one randomly chosen donor
 and partially digested with a combination of EcoRI and
 EcoRI Methylase. Size selected DNA was cloned into the
 pBAC3.6 vector at EcoRI sites"
 BASE COUNT 224 a 156 c 168 g 204 t
 ORIGIN

Query Match 64.0%; Score 16; DB 17; Length 752;
 Best Local Similarity 100.0%; Pred. No. 99;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TCCAGGAGAACAGAA 18
 |||||
 Db 27 TCCAGGAGAACAGAA 42

RESULT 44

BG681831
 LOCUS 602629761F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4754414 5',
 DEFINITION mRNA sequence.
 ACCESSION BG681831
 VERSION BG681831.1 GI:13913228
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS NTH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-re@mail.nih.gov

Tissue Procurement: James Cleaver, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM10615 row: j column: 15
 High quality sequence stop: 745.
 Location/Qualifiers

FEATURES

source
 1..806
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4754414"
 /clone_lib="NCI_CGAP_Skn4"
 /lab_host="DH10B (T1 phage-resistant)"
 /tissue_type="squamous cell carcinoma"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.5kb. Library constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."
 BASE COUNT 197 a 230 c 204 g 175 t
 ORIGIN

Query Match 64.0%; Score 16; DB 12; Length 806;
 Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GTCCAGGAGAACAGAA 17
 |||||
 Db 534 GTCCAGGAGAACAGAA 549

RESULT 45

BG978854
 LOCUS 602149311F2 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:4308099 5',
 DEFINITION mRNA sequence.

ACCESSION BG978854
 VERSION BG978854.1 GI:12346069
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-re@mail.nih.gov

Tissue Procurement: ATCC/BDT/DTF

cDNA Library Preparation: CLONTECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1CM1179 row: f column: 04

High quality sequence stop: 659.

FEATURES

source
 Location/Qualifiers
 1..825

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4308099"

/clone_lib="NIH_MGC_62"

/tissue_type="melanotic melanoma, high MDR"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: skin; Vector: pBMR-LIB (Clontech); Site_1:
 SfiI (ggccctcgcc); Site_2: SfiI (ggccattggcc);

Double-stranded cDNA was prepared from cell line RNA. 5'

and 3' adaptors were used in cloning as follows: 5'

adaptor sequence: 5'-ATTCTAGAGCCGCGCGGACATG-dt(30)BN-3'

sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor

(where B = A, C, or G and N = A, C, G, or T). Average

insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies

contained inserts by PCR. This library was enriched for

full-length clones and was constructed by Clontech

Laboratories (Palo Alto, CA)."

BASE COUNT 164 a 171 c 209 g 281 t

ORIGIN

Query Match 64.0%; Score 16; DB 12; Length 825;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GCCAGGAGAACAGAAAT 19

|||||

Db 784 GCCAGGAGAACAGAAAT 769

Search completed: November 24, 2002, 15:01:49
 Job time : 1554.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2002, 12:19:18 : Search time 211.5 Seconds
(without alignments)
266.194 Million cell updates/sec

Title: US-09-784-423-125
Perfect score: 25
Sequence: 1 TGTGCGAGGACAGAAATTTACAG 25

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 112599159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_101002.*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	100.0	25	AAZ27825	PCR primer for hum
c 2	25	100.0	1000	AAZ27732	Human DNA marker c
c 3	16	64.0	442	ABA58891	Human foetal liver
c 4	16	64.0	442	AAK07051	Human brain expres
c 5	16	64.0	442	AAK32792	Human bone marrow
c 6	16	64.0	442	AAI38606	Probe #7292 used t
c 7	16	64.0	442	ABS07591	Human genome-deriv
c 8	16	64.0	585	AAZ21494	Human sensory neur
c 9	16	64.0	2640	ABA09725	Human bone marrow

c 10	64.0	22080	20	AAK06751	Salmonella enteric
c 11	64.0	134525	11	AAQ04525	Total base sequenc
c 12	64.0	2944528	24	ABA03041	Listeria monocytog
c 13	60.0	134	20	AAK51637	Human secreted pro
c 14	60.0	244	21	AAK36452	Human secreted pro
c 15	60.0	478	22	AAI92971	Human polynucleoti
c 16	60.0	535	21	AAZ56768	Human transmembran
c 17	60.0	569	24	ABK62291	Rat sequence diffe
c 18	60.0	591	23	ABL26995	Drosophila melanog
c 19	60.0	1302	23	AAK56010	Salmonella typhi D
c 20	60.0	2433	23	AAK56011	DNA encoding novel
c 21	60.0	2433	23	AAK56012	Drosophila melanog
c 22	60.0	2649	23	AAK56013	Human immune/haema
c 23	60.0	11821	22	AAK26994	Human musculoskele
c 24	60.0	13395	22	AAK74325	Human immune/haema
c 25	60.0	15535	22	AAK36490	Human musculoskele
c 26	60.0	26241	22	ABA16222	Human nervous syst
c 27	60.0	28720	19	AAV49655	Human SC3 DNA. Ho
c 28	60.0	65632	21	AAK81502	N. meningitidis pa
c 29	60.0	349980	21	AAK21544	Neisseria meningit
c 30	56.0	123	21	AAK31935	Human secreted pro
c 31	56.0	246	21	AAK43073	Human 5' EST isola
c 32	56.0	318	24	ABN24228	Human ORFX polynuc
c 33	56.0	350	23	ABV01847	Human prostate exp
c 34	56.0	351	23	ABV11016	Human prostate exp
c 35	56.0	353	23	ABV02282	Human prostate exp
c 36	56.0	365	23	ABV21450	Human prostate exp
c 37	56.0	365	23	ABV27268	Human prostate exp
c 38	56.0	388	21	AAK30937	Human colon cancer
c 39	56.0	396	23	ABV32164	Human prostate exp
c 40	56.0	396	23	ABV41096	Human prostate exp
c 41	56.0	408	23	ABV32596	Human prostate exp
c 42	56.0	408	23	ABV41521	Human prostate exp
c 43	56.0	435	23	ABV11451	Human prostate exp
c 44	56.0	439	24	ABN75835	Human ORF82 CDNA,
c 45	56.0	454	22	AAK60019	Human immune/haema

ALIGNMENTS

RESULT 1
AAZ27825
ID AAZ27825 standard; DNA; 25 BP.
XX
AC AAZ27825;
XX
XX
DT 23-DEC-1999 (first entry)
XX
DE PCR primer for human DNA marker clone SI32.
XX
DE
XX
KW Tandem repeat sequence; DNA isolation; intermediate tandem repeat;
KW ITR sequence; pentanucleotide tandem repeat; stutter artifact;
KW DNA typing; DNA profiling; linkage analysis; criminal justice;
KW paternity testing; animal lineage analysis; microsatellite loci;
KW polymorphism detection; PCR primer; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9940194-A1.
XX
XX 12-AUG-1999.
XX
XX
XX 04-FEB-1999; 99WO-US02345.
XX
XX 04-FEB-1998; 98US-0018584.
XX
XX (PROM-) PROMEGA CORP.
XX
XX Schumm JW, Bacher JW;
XX WPI; 1999-590696/50.
DR

XX Isolating DNA containing intermediate tandem repeat sequences, useful
PT in DNA profiling
XX
PS Claim 30; Page 21; 111pp; English.
XX
CC This sequence is a PCR primer for a human DNA marker clone used in the
CC method of the invention. The method is for isolating a fragment of DNA
CC containing an intermediate tandem repeat (ITR) sequence using
CC hybridization selection, and comprises: (a) providing several DNA
CC fragments, at least one of which contains an ITR sequence, a region of
CC the DNA fragment which contains at least one repeat unit consisting of a
CC sequence of five, six or seven bases repeated in tandem at least two
CC times; (b) providing a stationary support having at least one
CC oligonucleotide associated with it, where the oligonucleotide includes a
CC sequence of nucleotides which is complementary to a portion of the ITR
CC sequence; and (c) combining the DNA fragments with the support under
CC conditions where the DNA fragments including the DNA fragment containing
CC the ITR sequence hybridize to the support. The method is particularly
CC used to isolate DNA containing pentanucleotide tandem repeat sequences as
CC well as to detect target ITR DNA sequences having a low incidence of
CC stutter artifacts (no more than 2.4%). The method is useful in DNA
CC profiling for linkage analysis, criminal justice, paternity testing and
CC the lineage of horses, dogs and other prize animals. The invention
CC overcomes problems related to the use of microsatellite loci in DNA
CC profiling. The method can detect polymorphisms with a low incidence of
CC stutter artifacts, which has previously been a problem in interpreting
CC allelic content of loci. The development of markers based on larger
CC repeat units, enables easier separation of the fragments on larger
CC electrophoretic gels. This allows the simultaneous analysis of more loci.
XX
SQ Sequence 25 BP; 9 A; 5 C; 6 G; 5 T; 0 other;

Query Match 100.0%; Score 25; DB 20; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGCCAGGACCAAGAAATTACAG 25
|||||
Db 1 TGTGCCAGGACCAAGAAATTACAG 25

RESULT 2
AAZ27732/c
ID AAZ27732 standard; DNA; 1000 BP.
XX
AC AAZ27732;
XX
DT 23-DEC-1999 (first entry)
XX
DE Human DNA marker clone S132.
XX
KW Tandem repeat sequence; DNA isolation; intermediate tandem repeat;
KW ITR sequence; pentanucleotide tandem repeat; stutter artifact;
KW DNA typing; DNA profiling; linkage analysis; criminal justice;
KW paternity testing; animal lineage analysis; microsatellite loci;
KW polymorphism detection; ds.
XX
OS Homo sapiens.
XX
PN WO9940194-A1.
XX
PD 12-AUG-1999.
XX
PF 04-FEB-1999; 99WO-US02345.
XX
PR 04-FEB-1998; 98US-0018594.
XX
PA (PROM-) PROMEGA CORP.
XX
PI Schumm JW, Bacher JW;
XX

DR WPI; 1999-590696/50.
XX Isolating DNA containing intermediate tandem repeat sequences, useful
PT in DNA profiling
XX
PS Claim 22; Page 78; 111pp; English.
XX
CC This sequence represents a human DNA marker clone used in the method of
CC the invention. The method is for isolating a fragment of DNA containing
CC an intermediate tandem repeat (ITR) sequence using hybridization
CC selection, and comprises: (a) providing several DNA fragments, at least
CC one of which contains an ITR sequence, a region of the DNA fragment which
CC contains at least one repeat unit consisting of a sequence of five, six
CC or seven bases repeated in tandem at least two times; (b) providing a
CC stationary support having at least one oligonucleotide associated with
CC it, where the oligonucleotide includes a sequence of nucleotides which is
CC complementary to a portion of the ITR sequence; and (c) combining the DNA
CC fragments with the support under conditions where the DNA fragments
CC including the DNA fragment containing the ITR sequence hybridize to the
CC support. The method is particularly used to isolate DNA containing
CC pentanucleotide tandem repeat sequences as well as to detect target ITR
CC DNA sequences having a low incidence of stutter artifacts (no more than
CC 2.4%). The method is useful in DNA profiling for linkage analysis,
CC criminal justice, paternity testing and other forensic and medical uses.
CC DNA typing is also useful for confirming the lineage of horses, dogs and
CC other prize animals. The invention overcomes problems related to the use
CC of microsatellite loci in DNA profiling. The method can detect
CC polymorphisms with a low incidence of stutter artifacts, which has
CC previously been a problem in interpreting allelic content of loci. The
CC development of markers based on larger repeat units, enables easier
CC separation of the fragments on electrophoretic gels. This allows the
CC simultaneous analysis of more loci.
XX
SQ Sequence 1000 BP; 302 A; 240 C; 231 G; 226 T; 1 other;

Query Match 100.0%; Score 25; DB 20; Length 1000;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGCCAGGACCAAGAAATTACAG 25
|||||
Db 726 TGTGCCAGGACCAAGAAATTACAG 702

RESULT 3
ABA5891/c
ID ABA5891 standard; DNA; 442 BP.
XX
AC ABA5891;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #7196.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00669.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
XX Claim 1; SEQ ID NO 7196; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 442 BP; 123 A; 85 C; 98 G; 136 T; 0 other;

Query Match 64.0%; Score 16; DB 22; Length 442;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTCGCCAGGAACAGA 16
|||||
DB 226 TGTCGCCAGGAACAGA 211

RESULT 4
AAK07051/c
ID AAK07051 standard; DNA; 442 BP.
XX
AC AAK07051;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 7042.
XX
XX Homo sapiens.
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.
OS
PN WO200157275-A2.
PP 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
DR
XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
XX Example 4; SEQ ID NO: 7042; 650pp + Sequence Listing; English.
PS
XX The present invention provides a number of single exon nucleic acid

Db 226 TGTGCCAGGACACAGA 211

RESULT 6
AAI38606/c
ID AAI38606 standard; DNA: 442 BP.
XX AC AAI38606;
XX DT 17-OCT-2001 (first entry)
XX DE Probe #7292 used to measure gene expression in human placenta sample.
XX DE Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200157272-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00663.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488897/53.
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
XX Claim 25; SEQ ID No 7292; 654pp; English.
XX The present invention relates to single exon nucleic acid probes (SENPs).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX Sequence 442 BP; 123 A; 85 C; 98 G; 136 T; 0 other;
XX
Query Match 64.0%; Score 16; DB 22; Length 442;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGTGCCAGGACACAGA 16
|||||
Db 226 TGTGCCAGGACACAGA 211

RESULT 7
ABS07591/c
ID ABS07591 standard; DNA: 442 BP.
XX AC ABS07591;
XX DT 19-AUG-2002 (first entry)
XX DE Human genome-derived single exon probe from lung SEQ ID No 7582.
XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

KW tuberculous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease.
OS Homo sapiens.
XX WO200186003-A2.
XX PD 15-NOV-2001.
XX PF 30-JAN-2001; 2001WO-US00665.
XX PR 04-FEB-2000; 2000US-180312P.
XX PR 26-MAY-2000; 2000US-207456P.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-234687P.
XX PR 27-SEP-2000; 2000US-236359P.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2002-114183/15.
XX DR Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples -
XX Claim 1; SEQ ID No 7582; 634pp; English.
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12387 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of
XX probes; the novel set of probes which hybridise at high stringency to a
XX nucleic acid expressed in the human lung; measuring gene expression in a
XX sample derived from human lung, comprising (a) contacting the array with
XX a collection of detectably labeled nucleic acids derived from human lung
XX mRNA, and (b) measuring the label detectably bound to each probe of
XX the array; identifying exons in a eukaryotic genome, comprising
XX (a) algorithmically predicting at least one exon from genomic sequences
XX of the eukaryote; and (b) detecting specific hybridisation of detectably
XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX having a fragment identical to the predicted exon, the probe is included
XX in the above mentioned microarray; assigning exons to a single gene,
XX comprising (a) identifying exons from genomic sequence by the method
XX above and (b) measuring the expression of each of the exons in several
XX tissues and/or cell types using hybridisation to a single exon
XX microarrays having a probe with the exon, where a common pattern of
XX expression of the exons in the tissues and/or cell types indicates that
XX the exons should be assigned to a single gene; a peptide comprising one
XX of 12011 sequences, mentioned in the specification, or encoded by the
XX probes/open reading frames (ORF). The probes are used for gene
XX expression analysis, and for identifying exons in a gene, particularly
XX using human lung derived mRNA and for the study of lung diseases
XX such as asthma, lung cancer, chronic obstructive pulmonary disease
XX (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
XX fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
XX Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
XX haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
XX pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
XX pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
XX and hyaline membrane disease. The present sequence is a single exon
XX probe of the invention.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from WIPO at

```

CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 442 BP; 123 A; 85 C; 98 G; 136 T; 0 other;

Query Match      64.0%; Score 16; DB 24; Length 442;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTCCGAGGACACAGA 16
   |||||
Db 226 TGTCCGAGGACACAGA 211

RESULT 8
AAZ21494/C
ID AAZ21494 standard; cDNA; 585 BP.
XX
AC AAZ21494;
XX
DT 03-DEC-1999 (first entry)
XX
DE Human sensory neurone specific 2a nucleotide sequence fragment #14.
XX
KW Sensory neurone specific 2a; SNS-2a; sodium channel protein; pain;
KW voltage gated; hypersensitivity; ss.
XX
OS Homo sapiens.
XX
PN WO9947670-A1.
XX
PD 23-SEP-1999.
XX
PF 18-MAR-1999; 99WO-GB00838.
XX
PR 18-MAR-1998; 98GB-0005793.
XX
PA (GLAX ) GLAXO GROUP LTD.
XX
PI Grose DT, Hick CA, Tate SN;
XX
DR WPI; 1999-562112/47.
XX
PT Mammalian sodium channel protein for treating pain and hypersensitivity
XX
PS Claim 6; Page 67-68; 73pp; English.
XX
CC AA221481 to AA221495 represent fragments of the human sensory neurone
CC specific 2a (SNS-2a) nucleotide sequence. SNS-2a is a sodium channel
CC protein. SNS-2a can be used in a method for the identification of a
CC modulator of a sodium channel, and for assaying for compounds which
CC modulate sodium flux. The sodium channel modulators can be used in a
CC medicament for the treatment of pain or hypersensitivity.
XX
SQ Sequence 585 BP; 181 A; 117 C; 112 G; 175 T; 0 other;

Query Match      64.0%; Score 16; DB 20; Length 585;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 GGAACGAGAAATTAC 23
   |||||
Db 471 GGAACGAGAAATTAC 456

RESULT 9
ABA09725
ID ABA09725 standard; DNA; 2640 BP.
XX
AC ABA09725;
XX
DT 15-JAN-2002 (first entry)
XX
DE Human bone marrow expressed oligonucleotide SEQ ID NO: 234.

XX Human; bone marrow; cytostatic; antirheumatic; antiarthritic; vulnery;
XX antiinflammatory; antibacterial; immunosuppressive; vasotropic; cancer;
XX antiparkinsonian; neuroprotective; nootropic; haemostatic; osteopathic;
XX antitumor; fungicide; antidiabetic; antisthmatic; antiallergic;
XX immunostimulant; analgesic; cerebroprotective; antianaemic; infection;
XX nervous system disorder; autoimmune disorder; inflammation; allergy; ds.
XX
OS Homo sapiens.
XX
PN WO200174836-A1.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US10472.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PR 23-AUG-2000; 2000US-0649267.
XX
PR 30-NOV-2000; 2000US-250583P.
XX
PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Ford JE, Boyle BJ;
XX WPI; 2001-626375/72.
XX
XX New bone marrow-expressed nucleic acids and polypeptides, useful for
XX diagnosis, treatment of inflammatory, autoimmune, neurological, cancer
XX PT and increasing hematopoiesis, stem cell survival and bone growth and
XX remodeling
XX
PS Claim 1; Page 295-296; 380pp; English.
XX
CC The present invention relates to bone marrow expressed polynucleotides
CC and proteins. These sequences can be used in the treatment of
CC inflammatory conditions (eg arthritis, Crohn's disease), cancer, central
CC and peripheral nervous system diseases and neuropathies, such as
CC Alzheimer's, Parkinson's and Huntington's diseases, spinal cord
CC disorders, head trauma, cerebrovascular diseases, myeloid and lymphoid
CC cell disorders, platelet disorders, stem cell disorders, bone
CC degenerative disorders, autoimmune disorders, for example multiple
CC sclerosis, diabetes and arthritis, viral and bacterial infections,
CC allergies and blood coagulation disorders. The present sequence is a DNA
XX of the invention.
XX
SQ Sequence 2640 BP; 706 A; 631 C; 674 G; 629 T; 0 other;

Query Match      64.0%; Score 16; DB 22; Length 2640;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GCCAGAACCCAGAAAT 19
   |||||
Db 2496 GCCAGAACCCAGAAAT 2511

RESULT 10
AA06751
ID AA06751 standard; DNA; 22080 BP.
XX
XX
AC AA06751;
XX
DT 26-APR-1999 (first entry)
XX
DE Salmonella enterica serovar B O antigen gene cluster.
XX
KW O antigen; diagnosis; wxr gene; wbaV gene; transferase; polymerase;
KW flippase; diarrhoea; haemorrhagic colitis; ds.
XX
OS Salmonella enterica.
XX
XX
FH Key Location/Qualifiers

```


FT CDS 4099..5184
 FT /tag= a
 FT /product= TDP-rhamnose_pathway_protein"
 FT /note= "rmlB gene; encodes AAW88335"
 FT 5184..6083
 FT /tag= b
 FT /product= TDP-rhamnose_pathway_protein"
 FT /note= "rmlD gene; encodes AAW88336"
 FT 6131..7009
 FT /tag= c
 FT /product= TDP-rhamnose_pathway_protein"
 FT /note= "rmlA gene; encodes AAW88337"
 FT 7010..7561
 FT /tag= d
 FT /product= TDP-rhamnose_pathway_protein"
 FT /note= "rmlC gene; encodes AAW88338"
 FT 7567..8559
 FT /tag= e
 FT /product= CDP-adequose_pathway_protein"
 FT /note= "ddhd gene; encodes AAW88339"
 FT 8556..9329
 FT /tag= f
 FT /product= CDP-adequose_pathway_protein"
 FT /note= "ddhA gene; encodes AAW88340"
 FT 9334..10413
 FT /tag= g
 FT /product= CDP-adequose_pathway_protein"
 FT /note= "ddhB gene; encodes AAW88341"
 FT 10440..11753
 FT /tag= h
 FT /product= CDP-adequose_pathway_protein"
 FT /note= "ddhC gene; encodes AAW88342"
 FT 11781..12680
 FT /tag= i
 FT /product= CDP-adequose_pathway_protein"
 FT /note= "abe gene; encodes AAW88343"
 FT 12762..14054
 FT /tag= j
 FT /product= flippase
 FT /note= "wzx gene, this region is specifically
 FT claimed in Claim 9; encodes AAW88344"
 FT 14059..15060
 FT /tag= k
 FT /product= abequosyl_transferase
 FT /note= "wbav gene, this region is specifically
 FT claimed in Claim 9; ; encodes AAW88345"
 FT 15379..16440
 FT /tag= l
 FT /product= mannosyl_transferase
 FT /note= "wbaU gene; encodes AAW88346"
 FT 16441..17385
 FT /tag= m
 FT /product= rhamnosyl_transferase
 FT /note= "wbaN gene; encodes AAW88347"
 FT 17386..18825
 FT /tag= n
 FT /product= GDP-mannose_pathway_protein
 FT /note= "manC gene; encodes AAW88348"
 FT 18812..20245
 FT /tag= o
 FT /product= GDP-mannose_pathway_protein
 FT /note= "mabB gene; encodes AAW88349"
 FT 20317..21747
 FT /tag= h
 FT /product= galactosyl_transferase
 FT /note= "wbap gene; encodes AAW88350"
 FT W09850531-A1.
 FT 12-NOV-1998.
 FT 01-MAY-1998;
 FT 98WO-AU00315.
 FT XX

PR 22-JUL-1997; 97AU-0008162.
 PR 01-MAY-1997; 97AU-0006545.
 XX (UNSY) UNIV SYDNEY.
 PA XX
 PA XX
 PI Reeves PR, Wang L;
 DR WPI; 1999-059669/05.
 DR P-PDB; AAW88335, AAW88336, AAW88337, AAW88338, AAW88339, AAW88340,
 DR AAW88341, AAW88342, AAW88343, AAW88344, AAW88345, AAW88346, AAW88347,
 DR AAW88348, AAW88349, AAW88350.
 XX
 PT Nucleic acid molecules specific for bacterial polysaccharide
 PT antigens - useful for detecting specific strains in, e.g. food,
 PT faeces or patient samples
 XX
 PS Claim 9; Fig 10; 165pp; English.
 XX
 CC This is the nucleotide sequence of a gene cluster involved in
 CC the biosynthesis of the Salmonella enterica serotype B (strain
 CC LT2) O antigen. 16 Open reading frames within the gene cluster
 CC were identified on the basis of homology to known sequences. The
 CC genes encode proteins (see AAW88335-50) involved in the synthesis of
 CC sugars present in the polysaccharide or oligosaccharide units. The use
 CC of processing of polysaccharide antigen, and in the transport
 CC of nucleic acid molecules derived from assembly and transport
 CC genes, particularly wbd (transferase), wzx (flippase) and wzy
 CC (polymerase) genes, within O antigen gene clusters improves the
 CC specificity of methods for the detection and identification of O
 CC antigens, e.g. in testing food- or faecal-derived samples, or
 CC samples from patients. The O antigen is a major virulence factor
 CC of S. enterica.
 XX
 SQ Sequence 22080 BP; 6284 A; 4028 C; 5152 G; 6616 T; 0 other;
 Query Match 64.0%; Score 16; DB 20; Length 22080;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 AGGACACAGAAATTTA 22
 |||||
 Db 16641 AGGACACAGAAATTTA 16656
 RESULT 11
 AAQ04525/c
 ID AAQ04525 standard; DNA; 134525 BP.
 AC AAQ04525;
 XX
 DT 01-OCT-1990 (first entry)
 DE Total base sequence of rice plant chloroplast DNA.
 XX
 KW Chloroplast; true grass; rice plant; ss.
 XX
 OS Oryza sativa.
 XX
 PN JP02100682-A.
 XX
 PD 12-APR-1990.
 XX
 PF 07-OCT-1988; 88JP-0251967.
 XX
 PR 07-OCT-1988; 88JP-0251967.
 XX
 PA (MITK) MITSUI TOATSU CHEM INC.
 XX
 DR WPI; 1990-159709/21.
 XX
 PT Chloroplast DNA of true grasses - used to produce various
 PT DNA base sequences by decomposition of rice plant DNA.
 XX

PS Claim 1; Fig 1; 20pp; Japanese.
XX The sequence is that of the whole of rice chloroplast DNA.
XX
SQ Sequence 134525 BP; 41249 A; 26129 C; 26331 G; 40816 T; 0 other;
Query Match 64.0%; Score 16; DB 11; Length 134525;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGTCCGAGGACACAGA 16
|||||
Db 47499 TGTCCGAGGACACAGA 47484

RESULT 12
ABR03041
ID ABR03041 standard; DNA; 2944528 BP.
XX
AC ABA03041;
XX
DT 05-FEB-2002 (first entry)
XX
DE Listeria monocytogenes EGD-e genome sequence.
XX
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease; ds.
XX
OS Listeria monocytogenes.
XX
PN W0200177335-A2.
XX
PD 18-OCT-2001.
XX
PF 11-APR-2001; 2001WO-FR01118.
XX
PR 11-APR-2000; 2000FR-0004629.
XX
PA (INSP) INST PASTEUR.
XX
PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Deboux P;
PI Dussurget O, Chetoui F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Dommann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Madueno E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
PI Rose M, Voss H;
XX
WPI; 2002-010914/01.
XX
Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
and prevention of Listeria and related bacterial infections, and
related polypeptides -
XX
Claim 1; SEQ ID No 1; 192pp; French.
XX
The present sequence is the genome sequence of Listeria monocytogenes
EGD-e. This sequence and fragments of this sequence are useful for
selecting probes and primers for detecting genes in L. monocytogenes and
related organisms, and to study genetic polymorphisms and other genomes.
Proteins (ABB47297-ABB50149) expressed from the present sequence are
useful for raising specific antibodies, identification of L.
monocytogenes and related organisms, and for biosynthesis and
biodegradation, especially biosynthesis of Vitamin B12. This sequence and
proteins encoded by it are also useful for selecting compounds that
regulate gene expression and cell replication and modulate L.
monocytogenes-related diseases. In addition, this sequence and proteins
encoded by it are useful in pharmaceutical and vaccines compositions for
the treatment or prevention of infections by L. monocytogenes and related
organisms.
XX
Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2944528 BP; 914202 A; 563301 C; 555061 G; 911964 T; 0 other;
Query Match 64.0%; Score 16; DB 24; Length 2944528;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 AACCCAGAAATTTACAG 25
|||||
Db 226486 AACCCAGAAATTTACAG 226501

RESULT 13
AA51637
ID AA51637 standard; cDNA; 134 BP.
XX
AC AA51637;
XX
DT 21-JUN-1999 (first entry)
XX
DE Human secreted protein 5' EST SEQ ID NO:216.
XX
KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition; ds.
XX
OS Homo sapiens.
XX
PN W09506549-A2.
XX
PD 11-FEB-1999.
XX
PF 31-JUL-1998; 98WO-IB01231.
XX
PR 01-AUG-1997; 97US-0905279.
XX
PA (GEST) GENSET.
XX
PI Duclert A, Dumas Milne Edwards J, Lacroix B;
XX
WPI; 1999-153779/13.
XX
P-PSDB; AA12859.
XX
New nucleic acids encoding human secreted proteins - obtained from
cDNA libraries derived from testis, ovary, uterus and spleen tissue
XX
Claim 1; Page 308; 522pp; English.
XX
AA51459 to AA51691 represent 5' expressed sequence tags (ESTs) for
human secreted proteins, and encode the proteins given in AA12681 to
AA12913, respectively. The proteins given represent the signal peptide
and an N-terminal fragment of a secreted protein. The nucleic acid
sequences can be used for producing secreted human gene products. They
can also be used to develop products for diagnosis and therapy. The
proteins obtained may have cytokine activity, cell
proliferation/differentiation activity, haematopoiesis regulating
activity, tissue growth regulating activity, chemotactic/chemokinetic
regulating activity, chemotactic/chemokinetic activity, reproductive hormone
thrombolytic activity, receptor/ligand activity, anti-inflammatory
activity, tumour inhibition activity or other activities. The products
can be used in forensic, gene therapy and chromosome mapping procedures.
The sequences can also be used for obtaining corresponding promoter
sequences. The nucleic acids encoding the signal peptide can be used for
directing extracellular secretion of a polypeptide or the insertion of a
polypeptide into a membrane, or importing a polypeptide into a cell.
XX
Sequence 134 BP; 30 A; 25 C; 29 G; 48 T; 2 other;
SQ
Query Match 60.0%; Score 15; DB 20; Length 134;

Human; transmembrane protein; HTMPN; diagnosis; immunospecific; antiproliferative; neuroprotective; immune disorder

KW reproductive disorder; smooth muscle disorder; neurological disorder;
 KW gastrointestinal disorder; developmental disorder;
 KW cell proliferative disorder; ss.

XX Homo sapiens.
 XX WO9961471-A2.
 XX 02-DEC-1999.
 XX 28-MAY-1999; 99MO-US11904.
 XX 29-MAY-1998; 98US-0087260.
 PR 02-JUL-1998; 98US-0091674.
 PR 02-OCT-1998; 98US-0102954.
 PR 24-NOV-1998; 98US-0109869.
 XX (INCY-) INCYTE PHARM INC.
 XX Tang YF, Lal P, Hillman JL, Yue H, Guegler KJ, Corley NC;
 PI Bandman O, Patterson C, Gorgone GA, Kaser MR, Baughn MR;
 PI Au-Young J;
 XX WPI; 2000-072605/06.
 DR P-PSDB; AAY57947.
 XX

PT Proteins, polynucleotides, vectors, host cells and antibodies used to
 PT diagnose, treat or prevent immune, reproductive, smooth muscle,
 PT neurological, gastrointestinal, developmental and cell proliferative
 PT disorders -
 PS Claim 9; Page 223; 229pp; English.

XX AZ55698 to AA256776 encode AAY57877 to AAY57955 which represent human
 CC transmembrane proteins designated HTPN-1 to HTPN-79, respectively.
 CC The transmembrane protein have immunospecific, antiproliferative and
 CC neuroprotective activities. The human transmembrane proteins,
 CC polynucleotides encoding them and other compositions and methods from
 CC the present invention, can be used for the diagnosis, treatment or
 CC prevention of immune, reproductive, smooth muscle, neurological,
 CC gastrointestinal, developmental and cell proliferative disorders. The
 CC HTPN's can be used to treat or prevent disorders associated with a
 CC decreased expression or activity of HTPN.
 XX Sequence 535 BP; 181 A; 93 C; 104 G; 157 T; 0 other;
 SQ

Query Match 60.0%; Score 15; DB 21; Length 535;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AGACACCAAAATTT 21
 DB 2 AGAACCAAAATTT 16

RESULT 17
 ABK62291

ID ABK62291 standard; cDNA; 569 BP.

XX AC ABK62291;

XX 18-JUN-2002 (first entry)

DE Rat sequence differentially expressed in response to a hepatotoxin #198.

KW Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
 KW differential expression; centrilobular necrosis; steatosis.
 XX

OS Rattus norvegicus.

XX WO200210453-A2.

XX 07-FEB-2002.

XX 30-JUL-2001; 2001WO-US23872.
 PF 31-JUL-2000; 2000US-222040P.
 XX 02-NOV-2000; 2000US-244880P.
 PR 11-MAY-2001; 2001US-250029P.
 PR 15-MAY-2001; 2001US-290645P.
 PR 22-MAY-2001; 2001US-292336P.
 PR 06-JUN-2001; 2001US-295798P.
 PR 13-JUN-2001; 2001US-297457P.
 PR 19-JUN-2001; 2001US-298884P.
 PR 09-JUL-2001; 2001US-303459P.
 XX (GENE-) GENE LOGIC INC.

PA Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;

XX WPI; 2002-241625/29.

DR Predicting toxic effects of compounds or the progression of these toxic
 XX effects by determining the changes in gene expression in tissues or
 XX cells exposed to the toxin and comparing these to gene expression in
 XX unexposed tissues or cells -
 PS Claim 1; Seq ID No 198; 239pp; English.

XX The invention relates to methods for predicting toxic effects of
 CC compounds or the progression of these toxic effects by determining the
 CC global changes in gene expression in tissues or cells exposed to the
 CC toxin and comparing these to gene expression in unexposed tissues or
 CC cells. Also included are methods of predicting at least one toxic
 CC effect of a compound or progression of a toxic effect, preferably the
 CC hepatotoxicity of a compound, comprising detecting the level of
 CC expression in a tissue or cell sample exposed to the compound of two or
 CC more genes listed in the specification, where differential expression of
 CC the genes is indicative of at least one toxic effect or progression.
 CC The method can also be used to identify an agent which modulates the
 CC toxic response and predict cellular pathways that a compound modulates
 CC in a cell. The methods utilise a set of at least two probes (on a solid
 CC support in kit form), where each of the probes comprises a sequence that
 CC specifically hybridises to a gene listed in the specification, a computer
 CC system comprising a database containing information identifying the
 CC expression level in a tissue or cell sample exposed to a hepatotoxin of a
 CC set of genes comprising at least two genes listed in the specification,
 CC and a user interface to view the information used to present information,
 CC identifying the expression level in a tissue or cell of at least one gene
 CC listed in the specification. The method is useful for elucidating global
 CC changes in gene expression and for identifying toxicity markers in
 CC tissues or cell exposed to a known toxin. The genes may be used as
 CC toxicity markers in drug screening and toxicity assays. The genes and
 CC gene expression information may be used as diagnostic markers for the
 CC prediction or identification of the physiological state of tissue or cell
 CC sample that has been exposed to a compound or agent. Hepatotoxicity
 CC is characterised by centrilobular necrosis and steatosis. The present
 CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
 CC which is differentially expressed in response to a hepatotoxic agent.

XX Sequence 569 BP; 146 A; 141 C; 132 G; 150 T; 0 other;

Query Match 60.0%; Score 15; DB 24; Length 569;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GCCAGGACCAAGAAA 18
 DB 240 GCCAGGACCAAGAAA 254

RESULT 18

ID ABL26995 standard; DNA; 591 BP.

XX ABL26995;

DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 32458.
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ds.
 KW Drosophila melanogaster.
 XX OS WO200171042-A2.
 XX PN 27-SEP-2001.
 XX PD 23-MAR-2001; 2001WO-US09231.
 XX PF 23-MAR-2000; 2000US-191637P.
 XX PR 11-JUL-2000; 2000US-0614150.
 XX XX (PEKE) PE CORP NY.
 XX PA Venter JC, Adams M, Li PWD, Myers EW;
 XX PI WPI; 2001-656860/75.
 XX DR
 XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
 XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
 XX PT interactions.
 XX PS Claim 1; SEQ ID NO 32458; 21pp + Sequence Listing; English.
 XX XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABBS7737-ABBS72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 591 BP; 167 A; 134 C; 175 G; 115 T; 0 other;
 Query Match 60.0%; Score 15; DB 23; Length 591;
 Best Local Similarity 100.0%; Pred. NO. 56;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 GCCAGGACCCAGAAA 18
 Db |||||||||||||||
 501 GCCAGGACCCAGAAA 515
 RESULT 19
 AAS56010
 ID AAS56010 standard; DNA; 1302 BP.
 XX AC
 XX AAS56010;
 XX DT 13-FEB-2002 (first entry)
 XX DE Salmonella typhi DNA for cellular proliferation protein #43.
 XX KW Antisense; ds; prokaryotic cellular proliferation gene;
 KW antibiotic; antibacterial; drug design.
 XX OS Salmonella typhi.
 XX PN WO200170955-A2.
 XX PD 27-SEP-2001.
 XX XX

PF 21-MAR-2001; 2001WO-US09180.
 XX XX
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX XX
 PA (ELIT-) ELITRA PHARM INC.
 XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX WPI; 2001-611495/70.
 DR P-PSDB; RAU38151.
 XX XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX PS Claim 27; Seq ID No 9647; 51pp; English.
 XX CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence encodes an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1302 BP; 322 A; 291 C; 352 G; 337 T; 0 other;
 Query Match 60.0%; Score 15; DB 23; Length 1302;
 Best Local Similarity 100.0%; Pred. NO. 57;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 8 GGAACCCAGAAATTTA 22
 Db |||||||||||||||
 735 GGAACCCAGAAATTTA 749
 RESULT 20
 AAS90113
 ID AAS90113 standard; cDNA; 2433 BP.
 XX AC
 XX AAS90113;
 XX DT 13-FEB-2002 (first entry)
 XX DE DNA encoding novel human diagnostic protein #25917.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX OS Homo sapiens.
 XX PN WO200175067-A2.
 XX PD 11-OCT-2001.
 XX XX

```
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR P-PSDB; ABG25926.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 1; SEQ ID NO 25917; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2433 BP; 654 A; 605 C; 620 G; 554 T; 0 other;
Query Match 60.0%; Score 15; DB 23; Length 2433;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 GGACACGAGAAATTTA 22
DB 1066 GGAACACGAGAAATTTA 1080
RESULT 21
AAS91291
ID AAS91291 standard; cDNA; 2433 BP.
XX
AC AAS91291;
XX
XX 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #27095.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
XX WO200175067-A2.
PN
XX 11-OCT-2001.
PD
XX
XX 30-MAR-2001; 2001WO-US08631.
PF
XX
```

```
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR P-PSDB; ABG27104.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 1; SEQ ID NO 27095; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2433 BP; 654 A; 605 C; 620 G; 554 T; 0 other;
Query Match 60.0%; Score 15; DB 23; Length 2433;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 GGACACGAGAAATTTA 22
DB 1066 GGAACACGAGAAATTTA 1080
RESULT 22
ABL26994
ID ABL26994 standard; DNA; 2649 BP.
XX
AC ABL26994;
XX
XX 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 32455.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
XX WO200171042-A2.
PN
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US09231.
PF
XX
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
```

```

XX PA (PEXE ) PE CORP NY.
XX PI Venter JC, Adams M, Li FWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions -
XX PS Claim 1; SEQ ID NO 32455; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins
XX CC (ABB57737-ABB72072).
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 2649 BP; 657 A; 596 C; 627 G; 769 T; 0 other;

Query Match 60.0%; Score 15; DB 23; Length 2649;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GCCAGGAGCCAGAAA 18
Db 1559 GCCAGGAGCCAGAAA 1573
|||||
|||||

RESULT 23
AAL36492
ID AAL36492 standard; DNA; 11821 BP.
AC AAL36492;
XX 08-JAN-2002 (first entry)
XX Human musculoskeletal system related polynucleotide SEQ ID NO 2857.
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anti-ulcer;
XX vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein;
XX musculoskeletal system; ds.
XX OS Homo sapiens.
XX PN W0200155367-A1.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US01338.
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216647.
XX 07-JUL-2000; 2000US-0216880.
XX 11-JUL-2000; 2000US-0217487.
XX 14-JUL-2000; 2000US-0217496.
XX 26-JUL-2000; 2000US-0218290.
XX 26-JUL-2000; 2000US-0220963.
XX 14-AUG-2000; 2000US-0220964.
XX 14-AUG-2000; 2000US-0224518.
XX 14-AUG-2000; 2000US-0224519.
XX 14-AUG-2000; 2000US-0225213.
XX 14-AUG-2000; 2000US-0225214.
XX 14-AUG-2000; 2000US-0225266.
XX 14-AUG-2000; 2000US-0225267.
XX 14-AUG-2000; 2000US-0225268.
XX 14-AUG-2000; 2000US-0225270.
XX 14-AUG-2000; 2000US-0225447.
XX 14-AUG-2000; 2000US-0225757.
XX 14-AUG-2000; 2000US-0225758.
XX 18-AUG-2000; 2000US-0225759.
XX 22-AUG-2000; 2000US-0226279.
XX 22-AUG-2000; 2000US-0226681.
XX 22-AUG-2000; 2000US-0226868.
XX 23-AUG-2000; 2000US-0227182.
XX 30-AUG-2000; 2000US-0228924.
XX 01-SEP-2000; 2000US-0229287.
XX 01-SEP-2000; 2000US-0229343.
XX 01-SEP-2000; 2000US-0229344.
XX 05-SEP-2000; 2000US-0229345.
XX 05-SEP-2000; 2000US-0229509.
XX 06-SEP-2000; 2000US-0229513.
XX 06-SEP-2000; 2000US-0230437.
XX 08-SEP-2000; 2000US-0230438.
XX 08-SEP-2000; 2000US-0231242.
XX 08-SEP-2000; 2000US-0231243.
XX 08-SEP-2000; 2000US-0231244.
XX 08-SEP-2000; 2000US-0231413.
XX 08-SEP-2000; 2000US-0231414.
XX 08-SEP-2000; 2000US-0232080.
XX 08-SEP-2000; 2000US-0232081.
XX 12-SEP-2000; 2000US-0232081.
XX 14-SEP-2000; 2000US-0232397.
XX 14-SEP-2000; 2000US-0232398.
XX 14-SEP-2000; 2000US-0232399.
XX 14-SEP-2000; 2000US-0232400.
XX 14-SEP-2000; 2000US-0232401.
XX 14-SEP-2000; 2000US-0233063.
XX 14-SEP-2000; 2000US-0233064.
XX 21-SEP-2000; 2000US-0233065.
XX 21-SEP-2000; 2000US-0234223.
XX 25-SEP-2000; 2000US-0234274.
XX 25-SEP-2000; 2000US-0234997.
XX 25-SEP-2000; 2000US-0234998.
XX 26-SEP-2000; 2000US-0235484.
XX 27-SEP-2000; 2000US-0235834.
XX 27-SEP-2000; 2000US-0235836.
XX 29-SEP-2000; 2000US-0236327.
XX 29-SEP-2000; 2000US-0236367.
XX 29-SEP-2000; 2000US-0236368.
XX 29-SEP-2000; 2000US-0236369.
XX 29-SEP-2000; 2000US-0236370.
XX 02-OCT-2000; 2000US-0236802.
XX 02-OCT-2000; 2000US-0237037.
XX 02-OCT-2000; 2000US-0237038.
XX 02-OCT-2000; 2000US-0237039.
XX 13-OCT-2000; 2000US-0237040.
XX 13-OCT-2000; 2000US-0239935.
XX 13-OCT-2000; 2000US-0239937.
XX 20-OCT-2000; 2000US-0240960.
XX 20-OCT-2000; 2000US-0241221.
XX 20-OCT-2000; 2000US-0241785.
XX 20-OCT-2000; 2000US-0241786.
XX 20-OCT-2000; 2000US-0241787.
XX 20-OCT-2000; 2000US-0241808.

```

```
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PR (HUMA-) HUMAN GENOME SCI INC.
XX
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI: 2001-451937/48.
XX
XX Isolated polypeptide for treating, preventing and/or prognosing
XX disorders related to the musculoskeletal system including
XX musculoskeletal cancers and also for testing and detection e.g.
XX diagnosis -
XX
XX Example 2; SEQ ID NO 2857; 781pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AAL34669-AAL37666) and proteins
XX (ABB03087-ABB04109) associated with the musculoskeletal system useful
XX for preventing, treating or ameliorating medical conditions e.g. by
XX protein or gene therapy. The genes are isolated from a range of human
XX tissues disclosed in the specification. The nucleic acids, proteins,
XX antibodies and (ant)agonists are useful in the diagnosis, treatment
XX and prevention of: (a) cancer, e.g. breast and ovarian cancer and
XX other cancers of the adrenal gland, bone, bone marrow, breast,
```

```
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anemia, autoimmune rheumatoid arthritis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC (c) cardiovascular disorders such as myocardial ischaemias; (d) wound
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and
CC parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pot_sequences.
XX
```

SO Sequence 11821 BP; 3068 A; 2740 C; 2786 G; 3227 T; 0 other;

Query Match 60.0%; Score 15; DB 22; Length 11821;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGCCAGGAAACACG 15
|||||
DB 3694 TGTGCCAGGAAACACG 3708

RESULT 24
AAK74325
ID AAK74325 standard; DNA; 13395 BP.

XX AAK74325;

AC AAK74325;

DT 07-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:29137.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.

OS Homo sapiens.

XX WO200157182-A2.

PN 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0215135.

PR 30-JUN-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216980.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.


```
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.

PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

XX (HUMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX
XX Disclosure; SEQ ID NO 29137; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of hematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK62169
XX represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 13395 BP; 3070 A; 3741 C; 3243 G; 3340 T; 1 other;
Query Match 60.0%; Score 15; DB 22; Length 13395;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGTGCCAGGAACACG 15
|||||
```

Db 3698 TGTGCCAGGACCAG 3712

RESULT 25

AAL36490

ID AAL36490 standard; DNA; 15535 BP.

XX

AC AAL36490;

XX

DT 08-JAN-2002 (first entry)

XX

DE Human musculoskeletal system related polynucleotide SEQ ID NO 2855.

XX

KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;

KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;

KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;

KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;

KW neurological disease; infection; human; secreted protein;

KW musculoskeletal system; ds.

XX

OS Homo sapiens.

XX

PN WO200155367-A1.

XX

PD 02-AUG-2001.

XX

PF 17-JAN-2001; 2001WO-US01338.

XX

PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214986.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216980.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226868.

PR 23-AUG-2000; 2000US-0227182.

PR 30-AUG-2000; 2000US-0227009.

PR 01-SEP-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.

PR 08-SEP-2000; 2000US-0232080.

PR 08-SEP-2000; 2000US-0232081.

PR 12-SEP-2000; 2000US-0231968.

PR 14-SEP-2000; 2000US-0232397.

PR 14-SEP-2000; 2000US-0232398.

PR 14-SEP-2000; 2000US-0232399.

PR 14-SEP-2000; 2000US-0232400.

PR 14-SEP-2000; 2000US-0232401.

PR 14-SEP-2000; 2000US-0233063.

PR 14-SEP-2000; 2000US-0233064.

PR 21-SEP-2000; 2000US-0233065.

PR 21-SEP-2000; 2000US-0234223.

PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.

PR 25-SEP-2000; 2000US-0234998.

PR 26-SEP-2000; 2000US-0235484.

PR 27-SEP-2000; 2000US-0235834.

PR 27-SEP-2000; 2000US-0235836.

PR 29-SEP-2000; 2000US-0236327.

PR 29-SEP-2000; 2000US-0236367.

PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236369.

PR 29-SEP-2000; 2000US-0236370.

PR 02-OCT-2000; 2000US-0236802.

PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.

PR 02-OCT-2000; 2000US-0237039.

PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.

PR 13-OCT-2000; 2000US-0239937.

PR 20-OCT-2000; 2000US-0240960.

PR 20-OCT-2000; 2000US-0241221.

PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241786.

PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241808.

PR 20-OCT-2000; 2000US-0241809.

PR 20-OCT-2000; 2000US-0241826.

PR 01-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0246474.

PR 08-NOV-2000; 2000US-0246475.

PR 08-NOV-2000; 2000US-0246476.

PR 08-NOV-2000; 2000US-0246477.

PR 08-NOV-2000; 2000US-0246478.

PR 08-NOV-2000; 2000US-0246523.

PR 08-NOV-2000; 2000US-0246524.

PR 08-NOV-2000; 2000US-0246525.

PR 08-NOV-2000; 2000US-0246526.

PR 08-NOV-2000; 2000US-0246527.

PR 08-NOV-2000; 2000US-0246528.

PR 08-NOV-2000; 2000US-0246532.

PR 08-NOV-2000; 2000US-0246609.

PR 08-NOV-2000; 2000US-0246610.

PR 08-NOV-2000; 2000US-0246611.

PR 08-NOV-2000; 2000US-0246613.

PR 17-NOV-2000; 2000US-0249207.

PR 17-NOV-2000; 2000US-0249208.

PR 17-NOV-2000; 2000US-0249209.

PR 17-NOV-2000; 2000US-0249210.

PR 17-NOV-2000; 2000US-0249211.

PR 17-NOV-2000; 2000US-0249212.

PR 17-NOV-2000; 2000US-0249213.

PR 17-NOV-2000; 2000US-0249214.

PR 17-NOV-2000; 2000US-0249215.

PR 17-NOV-2000; 2000US-0249216.

PR 17-NOV-2000; 2000US-0249217.

PR 17-NOV-2000; 2000US-0249218.

PR 17-NOV-2000; 2000US-0249244.

PR 17-NOV-2000; 2000US-0249245.

PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0259678.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 DR WPI; 2001-451937/48.
 XX
 PT Isolated polypeptide for treating, preventing and/or prognosing
 PT disorders related to the musculoskeletal system including
 PT musculoskeletal cancers and also for testing and detection e.g.
 PT diagnosis -
 XX
 PS Example 2; SEQ ID NO 2855; 781pp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (AAL34669-AAL37666) and proteins
 CC (AB03087-AB04109) associated with the musculoskeletal system useful
 CC for preventing, treating or ameliorating medical conditions e.g. by
 CC protein or gene therapy. The genes are isolated from a range of human
 CC tissues disclosed in the specification. The nucleic acids, proteins,
 CC antibodies and (ant)agonists are useful in the diagnosis, treatment
 CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
 CC other cancers of the adrenal gland, bone, bone marrow, breast,
 CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
 CC (c) cardiovascular disorders such as myocardial ischaemia; (d) wound
 CC healing; (e) neurological diseases such as cerebral anoxia and epilepsy;
 CC and (f) infectious diseases such as viral, bacterial, fungal and
 CC parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 15535 BP; 3779 A; 3966 C; 3788 G; 4002 T; 0 other;
 Query Match 60.0%; Score 15; DB 22; Length 15535;
 Best Local Similarity 100.0%; Pred. No. 61;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGTGTCAGGACACG 15
 Db 3783 TGTGTCAGGACACG 3797
 |||||
 RESULT 26
 ID ABA16222 standard; DNA; 26241 BP.
 XX
 AC ABA16222;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Human nervous system related polynucleotide SEQ ID NO 8553.
 XX

KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
 KW antiparkinsonian; antisticking; antianaemic; antiarthritic; cancer;
 KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
 KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
 KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200159063-A2.
 XX
 XX 16-AUG-2001.
 PD
 XX
 XX 17-JAN-2001; 2001WO-US01334.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205115.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225278.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0232081.
 PR 14-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.

PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.

PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251889.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2000US-0254097.
PR 05-JAN-2001; 2000US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-541565/60.
DR Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system
PT cancers and metastases -
XX Disclosure; SEQ ID NO 8553; 1701pp + Sequence Listing; English.
XX The invention relates to novel genes (AB11004-ABA21534) and proteins
CC (AB114678-AB118001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 26241 BP; 7468 A; 5260 C; 5964 G; 7549 T; 0 other;

Query Match 60.0%; Score 15; DB 22; Length 26241;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGAACAC 15
Db 10724 TGTGCCAGGAACAC 10738
|||||

RESULT 27
AAV49655
ID AAV49655 standard; DNA; 28720 BP.
XX AC AAV49655;
XX 23-OCT-1998 (first entry)
DT Human SC3 DNA.
DE
XX
KW Scavenger receptor cysteine rich domain; SRCR; diagnosis; treatment;
KW nervous system; medullo-blastoma; glioma; breast; detection; SC3;
KW autoantibody; ss.
OS Homo sapiens.
XX
XX WO9830687-A2.
PN 16-JUL-1998.
PD
XX
XX 09-JAN-1998; 98WO-DE00096.
PF 18-JUL-1997; 97DE-1030997.
XX 09-JAN-1997; 97DE-1000519.
PR

XX PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX PI Mollenhauer J, Poustka A;

XX XX WPI; 1998-399136/34.

XX DR Proteins containing scavenger receptor, cysteine rich domain -

XX PT useful for diagnosis and treatment of tumours

XX XX Disclosure; Fig 5c; 54pp; German.

XX CC This nucleotide sequence contains a fragment of a gene which is capable

CC of encoding a protein which contains a SRCR (scavenger receptor,

CC cysteine-rich) domain. The gene and encoded protein can be used to

CC diagnose or treat tumours, particularly of the nervous system

CC (medullo-blastoma or glioma) or breast. The DNA sequence and probes

CC derived from it, are used to identify genes that express SRCR-domain

CC containing proteins, to determine the form in which these proteins exist

CC and to assess the significance of individual forms on cellular

CC properties. The protein can be used to detect the presence of

CC autoantibodies and antibodies which regulate its expression.

XX SQ Sequence 28720 BP; 6956 A; 6728 C; 6982 G; 8054 T; 0 other;

Query Match 60.0%; Score 15; DB 19; Length 28720;

Best Local Similarity 100.0%; Pred. No. 62;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CAGGACCAAGAAATT 20

Db 5047 CAGGACCAAGAAATT 5061

|||||||

RESULT 28

AAA81502

ID AAA81502 standard; DNA; 65632 BP.

XX AC AAA81502;

XX DT 04-DEC-2000 (first entry)

XX DE N. meningitidis partial DNA sequence gnm_49 SEQ ID NO:49.

XX KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;

XX KW antigen; vaccine; diagnosis; infection; antibacterial; identification;

XX KW Meningococcus B; MenB; ds.

XX OS Neisseria meningitidis.

XX XX WO200022430-A2.

XX PN 20-APR-2000.

XX PD 08-OCT-1999; 99WO-US23573.

XX PF 09-OCT-1998; 98US-0103794.

XX PR 30-APR-1999; 99US-0132068.

XX XX (CHIR) CHIRON CORP.

XX PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;

PI Masiagnani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;

PI Rappuoli R, Pizza M;

XX XX WPI; 2000-318079/27.

XX DR Isolated nucleotide sequences of Neisseria meningitidis which can be

XX PT used in the diagnosis and treatment of N. meningitidis infection and

PT other Neisserial infections, for example, N.gonorrhoea -

XX PS Claim 7; Page 1331-1350; 1760pp; English.

CC The present invention describes methods of obtaining immunogenic

CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414

CC represent specifically claimed Neisseria meningitidis genomic DNA

CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent

CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to

CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the

CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to

CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF

CC sequences, which are all used in the exemplification of the present

CC invention. The nucleic acid sequences, protein sequences, and antibodies

CC against them, can be used in the manufacture of a composition. The

CC composition can be used as a medicament (or in the manufacture of a

CC medicament) for treating, preventing or diagnosing infection due to

CC Neisserial bacteria. For example, some of the identified proteins could

CC be components of vaccines against Meningococcus B; against all serotypes;

CC and/or against all pathogenic Neisseriae. Identification of sequences

CC from the bacterium will also facilitate production of biological probes,

CC particularly organism-specific probes. Attempts to make efficacious

CC Meningococcus B vaccines have failed mainly due to antigen tolerance.

CC Multivalent vaccines have also been tried but none have successfully

CC overcome antigenic variability. The provision of further, complete

CC sequences may provide an opportunity to identify secreted or surface

CC exposed proteins that may be presumed targets for the immune system and

CC which are not antigenically variable or at least more conserved than

CC other more variable regions.

XX SQ Sequence 65632 BP; 16704 A; 16327 C; 15474 G; 17126 T; 1 other;

Query Match 60.0%; Score 15; DB 21; Length 65632;

Best Local Similarity 100.0%; Pred. No. 64;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 GAACGACAGAAATTTAC 23

|||||||

Db 37275 GAACGACAGAAATTTAC 37289

RESULT 29

AAF21544/c

ID AAF21544 standard; DNA; 349980 BP.

XX AC AAF21544;

XX DT 13-MAR-2001 (first entry)

XX DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:1.

XX KW Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;

XX KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;

XX KW ds.

XX OS Neisseria meningitidis.

XX PN WO2000066791-A1.

XX PD 09-NOV-2000.

XX PF 08-MAR-2000; 2000WO-US05928.

XX PR 30-APR-1999; 99US-0132068.

XX PR 08-OCT-1999; 99WO-US23573.

XX PR 28-FEB-2000; 2000GB-0004695.

XX XX (CHIR) CHIRON CORP.

XX PA (GENO-) INST GENOMIC RES.

XX PI Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Masiagnani V;

PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;

PI Frazer CM, Grandi G;

XX DR WPI; 2000-647603/62.

XX PT Neisseria meningitidis B full length genome sequence and open reading

PT frames are used to detect, treat and prevent Neisserial infections -
 PS Claim 7; Appendix A; 692pp; English.
 XX
 CC The present invention describes the full length genome of
 CC Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
 CC to AAF21613 represent fragments of the NMB genomic sequence, as the
 CC sequence was too long to go in a record on its own it was split into 8
 CC sequences which overlap each other at the beginning and end of each
 CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
 CC the beginning of AAF21607, the last 49980 bp of AAF21607 is repeated at
 CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
 CC Neisseria proteins given in AAF21589 to AAF21593, and AAF21594 to
 CC AAF21606 represent PCR primers which are used in the exemplification of
 CC the present invention. The NMB genome and fragments from it have
 CC antibacterial activity, and can be used in vaccines and gene therapy.
 CC Neisseria nucleic acids, proteins and/or antibodies which binds to the
 CC proteins can be used in compositions for treating or preventing infection
 CC due to Neisserial bacteria or as a diagnostic reagent for detecting the
 CC presence of Neisserial bacteria or of antibodies raised to Neisserial
 CC bacteria. Computers, computer memory, computer storage medium or computer
 CC databases can be used in a search to identify open reading frames (ORFs)
 CC or coding sequences within the NMB genome. The DNA sequences provide
 CC further opportunities to find antigenic or immunogenic proteins which are
 CC more effective in vaccines than the outer membrane proteins currently
 CC used.
 XX
 SQ Sequence 349980 BP; 83241 A; 85091 C; 95206 G; 86442 T; 0 other;
 Query Match 60.0%; Score 15; DB 21; Length 349980;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 GAACGAGAAATTAC 23
 Db 96883 GAACGAGAAATTAC 96869
 |||||
 RESULT 30
 AAC1935
 ID AAC1935 standard; cDNA; 123 BP.
 AC AAC1935;
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein 5' EST, SEQ ID NO: 36010.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-0200610.
 XX
 PR 26-FEB-1999; 99US-0122487.
 XX
 PA (GEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 DR WPI; 2000-500381/45.
 XX
 CC New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 1; SEQ ID 36010; 71pp + CD-ROM; English.
 XX

CC The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.
 XX
 SQ Sequence 123 BP; 44 A; 24 C; 17 G; 38 T; 0 other;
 Query Match 56.0%; Score 14; DB 21; Length 123;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 11 ACCAGAAATTAC 24
 Db 32 ACCAGAAATTAC 45
 |||||
 RESULT 31
 AAZ43073/c
 ID AAZ43073 standard; cDNA; 246 BP.
 XX
 AC AAZ43073;
 XX
 DT 01-FEB-2000 (first entry)
 XX
 DE Human 5' EST isolated from a cDNA library SEQ ID NO:1620.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;
 KW gene therapy; chromosome mapping; upstream regulatory sequence;
 KW forensic; location; development; protein synthesis; stability;
 KW regulation; identification; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9953051-A2.
 XX
 PD 21-OCT-1999.
 XX
 PF 09-APR-1999; 99WO-IB00712.
 XX
 PR 09-APR-1998; 98US-0057719.
 XX
 PR 28-APR-1998; 98US-0069047.
 XX
 PA (GEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 DR WPI; 2000-038446/03.
 XX
 PT Novel secreted protein 5' expressed sequence tag sequences used in
 PT diagnostic, forensic, gene therapy, and chromosome mapping procedures
 XX
 PS Claim 1; Page 836; 837pp; English.
 XX
 CC AAZ42265 to AAZ43075 represent novel 5' expressed sequence tag (EST)
 CC sequences, corresponding to human secreted proteins. AAY64651 to
 CC AAY6438 represent the EST-related proteins corresponding to AAZ42265 to
 CC AAZ43052. The 5' ESTs can be used for producing secreted human gene
 CC products. They can be used to identify and isolate 5' untranslated
 CC regions (UTRs) and upstream regulatory regions which control the
 CC location, development stage, rate, and quantity of protein synthesis, as
 CC well as stability of mRNA. The ESTs are also useful as probes for
 CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can
 CC also be used in forensic procedures to identify individuals, or in

CC diagnostic procedures to identify individuals having genetic diseases
 CC resulting from abnormal gene expression. The products may also be used in
 CC gene therapy protocols. The nucleic acids encoding signal peptides can be
 CC used for directing extracellular secretion of a polypeptide or the
 CC insertion of a polypeptide into a membrane, or importing a polypeptide
 CC into a cell. The proteins encoded by the EST sequences may be useful in
 CC treating a variety of human conditions. Secreted proteins have
 CC therapeutic value, and the identification of new secreted proteins is
 CC valuable. AA242249 to AA242264 and AA242264 to AA242265 represent
 CC sequences used in the exemplification of the present invention.

XX Sequence 246 BP; 64 A; 40 C; 38 G; 102 T; 2 other;

Query Match 56.0%; Score 14; DB 21; Length 246;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGAACCA 14
 |||||
 DB 175 TGTGCCAGGAACCA 162

RESULT 32
 ABN24228/c
 ID ABN24228 standard; cDNA; 318 BP.

AC ABN24228;

DT 24-JUN-2002 (first entry)

DE Human ORFX polynucleotide sequence SEQ ID NO:16933.

XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis; gene; ss.

OS Homo sapiens.

XX WO200192523-A2.

PN 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US10836.

XX 30-MAY-2000; 2000US-206132P.

XX 29-AUG-2000; 2000US-228716P.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach MD;

PI WPI; 2002-106308/14.

DR P-PSDB; ABP08476.

XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders

XX Disclosure; SEQ ID 16933; 1037pp; English.

XX The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification). ABN15762 to ABN27252 encode the human ORFX
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the

CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative diseases, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 318 BP; 41 A; 103 C; 112 G; 62 T; 0 other;

Query Match 56.0%; Score 14; DB 24; Length 318;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GCCAGGAACCAAGAA 17

|||||

DB 195 GCCAGGAACCAAGAA 182

RESULT 33
 ABV01847/c

XX ID ABV01847 standard; cDNA; 350 BP.

AC ABV01847;

XX DT 13-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 1838.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US05171.

XX PR 17-FEB-2000; 2000US-183319P.

XX PR 16-MAR-2000; 2000US-189862P.

XX PR 25-MAY-2000; 2000US-207454P.

XX PR 09-JUN-2000; 2000US-211314P.

XX PR 18-JUL-2000; 2000US-219007P.

XX PR 13-DEC-2000; 2000US-255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer

XX Claim 1; Page 378; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;

CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX
 SQ Sequence 350 BP; 76 A; 71 C; 81 G; 122 T; 0 other;
 Query Match 56.0%; Score 14; DB 23; Length 350;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 11 ACCAGAAATTAC 24
 DB 311 ACCAGAAATTAC 298
 RESULT 34
 ABV11016/c
 ID ABV11016 standard; cDNA; 351 BP.
 XX
 AC ABV11016;
 DT 13-SEP-2002 (first entry)
 XX Human prostate expression marker cDNA 11007.
 DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX Homo sapiens.
 OS
 XX WO200160860-A2.
 PN 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US05171.
 XX
 PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 DR WPI; 2001-662795/76.
 XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX
 PS Claim 1; Page 1783; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX
 SQ Sequence 351 BP; 76 A; 71 C; 81 G; 123 T; 0 other;
 Query Match 56.0%; Score 14; DB 23; Length 351;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 11 ACCAGAAATTAC 24
 DB 311 ACCAGAAATTAC 298
 RESULT 35
 ABV02282/c
 ID ABV02282 standard; cDNA; 353 BP.
 XX
 AC ABV02282;
 DT 13-SEP-2002 (first entry)
 XX Human prostate expression marker cDNA 2273.
 DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX Homo sapiens.
 OS
 XX WO200160860-A2.
 PN 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US05171.
 XX
 PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 DR WPI; 2001-662795/76.
 XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX
 PS Claim 1; Page 441; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

SQ Sequence 353 BP; 89 A; 83 C; 73 G; 108 T; 0 other;
 Query Match 56.0%; Score 14; DB 23; Length 353;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGAACCA 14
 ID ABV21450/c
 XX 259 TGTGCCAGGAACCA 246

RESULT 36
 ABV21450/c
 ID ABV21450 standard; CDNA; 365 BP.
 XX AC ABV21450;
 XX DT 13-SEP-2002 (first entry)
 XX DE Human prostate expression marker cDNA 21441.
 XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 XX KW pharmacogenomic marker; gene; ss.
 XX OS Homo sapiens.
 XX PN WO200160860-A2.
 XX PD 23-AUG-2001.
 XX PF 20-FEB-2001; 2001WO-US05171.
 XX PR 17-FEB-2000; 2000US-183319P.
 XX PR 16-MAR-2000; 2000US-189862P.
 XX PR 25-MAY-2000; 2000US-207454P.
 XX PR 09-JUN-2000; 2000US-211314P.
 XX PR 18-JUL-2000; 2000US-219007P.
 XX PR 13-DEC-2000; 2000US-255281P.
 XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX PI Schlegel R, Endege WO, Monahan JE;
 XX DR WPI; 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer -
 Claim 1; Page 3576; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for:
 (a) assessing whether a patient is afflicted with prostate cancer;
 (b) monitoring the progression of prostate cancer in a patient;
 (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;
 (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;
 (e) selecting a composition for inhibiting prostate cancer in a patient;
 (f) assessing the prostate cell carcinogenic potential of a compound;
 (g) determining whether prostate cancer has metastasized in a patient;
 (h) assessing the aggressiveness or indolence of prostate cancer in a patient;
 (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

SQ Sequence 365 BP; 92 A; 78 C; 69 G; 123 T; 3 other;
 Query Match 56.0%; Score 14; DB 23; Length 365;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGAACCA 14
 ID ABV27268/c
 XX 210 TGTGCCAGGAACCA 197

RESULT 37
 ABV27268/c
 ID ABV27268 standard; CDNA; 365 BP.
 XX AC ABV27268;
 XX DT 16-SEP-2002 (first entry)
 XX DE Human prostate expression marker cDNA 27259.
 XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 XX KW pharmacogenomic marker; gene; ss.
 XX OS Homo sapiens.
 XX PN WO200160860-A2.
 XX PD 23-AUG-2001.
 XX PF 20-FEB-2001; 2001WO-US05171.
 XX PR 17-FEB-2000; 2000US-183319P.
 XX PR 16-MAR-2000; 2000US-189862P.
 XX PR 25-MAY-2000; 2000US-207454P.
 XX PR 09-JUN-2000; 2000US-211314P.
 XX PR 18-JUL-2000; 2000US-219007P.
 XX PR 13-DEC-2000; 2000US-255281P.
 XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX PI Schlegel R, Endege WO, Monahan JE;
 XX DR WPI; 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer -
 Claim 1; Page 5536; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for:
 (a) assessing whether a patient is afflicted with prostate cancer;
 (b) monitoring the progression of prostate cancer in a patient;
 (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;
 (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;
 (e) selecting a composition for inhibiting prostate cancer in a patient;
 (f) assessing the prostate cell carcinogenic potential of a compound;
 (g) determining whether prostate cancer has metastasized in a patient;
 (h) assessing the aggressiveness or indolence of prostate cancer in a patient;
 (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

SQ Sequence 365 BP; 92 A; 78 C; 69 G; 123 T; 3 other;
 Query Match 56.0%; Score 14; DB 23; Length 365;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGAACCA 14
 ID ABV27268/c
 XX 210 TGTGCCAGGAACCA 197

RESULT 38

AAH30937/c
 ID AAH30937 standard; cDNA; 388 BP.
 AC AAH30937;
 XX
 DT 27-JUL-2001 (first entry)
 XX
 DE Human colon cancer cell line Kml2L4-A cDNA library derived sequence #871.
 XX
 KW Human; diagnosis; colon cancer; cancer; malignant; chromosome mapping;
 KW detection; colon cancer cell line Kml2L4-A; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200018916-A2.
 XX
 PD 06-APR-2000.
 XX
 PF 23-SEP-1999; 99WO-US22226.
 XX
 PR 28-SEP-1998; 98US-0102161.
 PR 28-SEP-1998; 98US-0102180.
 PR 29-SEP-1998; 98US-0102380.
 PR 08-OCT-1998; 98US-0103815.
 PR 27-OCT-1998; 98US-0105877.
 XX
 PA (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.
 XX
 PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
 PI Reinhard C, Glese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
 PI Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
 PI Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Grain B;
 XX
 DR WPI; 2000-293155/25.
 XX
 PT Polynucleotide library comprising 1079 defined sequences, useful in
 PT the form of an array to detect cancer or susceptibility to cancer -
 XX
 PS Claim 1; Page 433; 502pp; English.
 XX
 CC The present invention describes a library of polynucleotides comprising
 CC 1079 nucleotide sequences (given in AAH30067 to AAH31145). Also described
 CC are: (1) an isolated polynucleotide (I) having at least 90% identity to
 CC one of the 1079 sequences; (2) a recombinant host cell containing (1);
 CC (3) an isolated polypeptide (II) encoded by (1); (4) an antibody that
 CC specifically binds to (II); (5) a vector comprising (1); and (6) a method
 CC of detecting differentially expressed genes correlated with a cancerous
 CC state of a mammalian cell comprising detecting a gene product encoded by
 CC 65 of the 1079 sequences given in the specification. The polynucleotides
 CC are used to monitor patients having (or susceptible) to cancer to detect
 CC potentially malignant events at a molecular level before they are
 CC detectable at a gross morphological level. The polynucleotides are also
 CC useful for monitoring the efficacy of various therapies and preventive
 CC interventions. Polynucleotide probes based on the disclosed sequences
 CC are useful for chromosome mapping and detection of transcription levels.
 CC The 1079 polynucleotide sequences were derived from a human colon cancer
 CC cell line Kml2L4-A cDNA library.
 XX
 SQ Sequence 388 BP; 87 A; 97 C; 51 G; 153 T; 0 other;
 Query Match 56.0%; Score 14; DB 21; Length 388;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 8 GGAACCAAGAAATTT 21
 DB 370 GGAACCAAGAAATTT 357
 RESULT 39
 ABV32164/c
 ID ABV32164 standard; cDNA; 396 BP.

XX
 AC ABV32164;
 XX
 DT 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 32155.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US05171.
 XX
 PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 DR WPI; 2001-662795/76.
 XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX
 PS Claim 1; Page 6877; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX
 SQ Sequence 396 BP; 85 A; 85 C; 97 G; 129 T; 0 other;
 Query Match 56.0%; Score 14; DB 23; Length 396;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 11 ACCAGAAATTTACA 24
 DB 357 ACCAGAAATTTACA 344
 RESULT 40
 ABV41096/c
 ID ABV41096 standard; cDNA; 396 BP.
 XX
 AC ABV41096;
 XX
 DT 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 41087.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX Homo sapiens.
 CS
 XX WO200160860-A2.
 PN 23-AUG-2001.
 XX
 XX 20-FEB-2001; 2001WO-US05171.
 XX 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PA Schlegel R, Endege WO, Monahan JE;
 XX WPI; 2001-662795/76.
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 PS Claim 1; Page 8268; 11750pp; English.
 XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX Sequence 396 BP; 85 A; 85 C; 97 G; 129 T; 0 other;
 SQ
 Query Match 56.0%; Score 14; DB 23; Length 396;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 11 ACCAGAAATTAC 24
 Db 357 ACCAGAAATTAC 344
 RESULT 41
 ABV32596/C
 ID ABV32596 standard; cDNA; 408 BP.
 XX
 AC ABV32596;
 XX
 XX 16-SEP-2002 (first entry)
 DT Human prostate expression marker cDNA 32587.
 DE
 XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX Homo sapiens.
 OS
 XX

PN WO200160860-A2.
 XX 23-AUG-2001.
 PD
 XX 20-FEB-2001; 2001WO-US05171.
 XX 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PA Schlegel R, Endege WO, Monahan JE;
 XX WPI; 2001-662795/76.
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 PS Claim 1; Page 6948; 11750pp; English.
 XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX Sequence 408 BP; 98 A; 92 C; 89 G; 129 T; 0 other;
 SQ
 Query Match 56.0%; Score 14; DB 23; Length 408;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGTGCCAGGAACCA 14
 Db 253 TGTGCCAGGAACCA 240
 RESULT 42
 ABV41521/C
 ID ABV41521 standard; cDNA; 408 BP.
 XX
 AC ABV41521;
 XX
 XX 16-SEP-2002 (first entry)
 DT Human prostate expression marker cDNA 41512.
 DE
 XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX Homo sapiens.
 OS
 XX WO200160860-A2.
 PN 23-AUG-2001.
 PD
 XX 20-FEB-2001; 2001WO-US05171.
 PF
 XX

PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 XX Schlegel R, Endege WO, Monahan JE;
 PI
 DR WPI; 2001-662795/76.
 XX
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer
 XX
 PS Claim 1; Page 8341; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX
 SQ Sequence 408 BP; 98 A; 92 C; 89 G; 129 T; 0 other;
 Query Match 56.0%; Score 14; DB 23; Length 408;
 Best Local Similarity 100.0%; Pred. NO. 2e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGTGCCAGGAACCA 14
 Db 253 TGTGCCAGGAACCA 240
 RESULT 43
 ABV11451/c
 ID ABV11451 standard; cDNA; 435 BP.
 XX
 AC ABV11451;
 XX
 DT 13-SEP-2002 (first entry)
 DE
 DE Human prostate expression marker cDNA 11442.
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 KW Homo sapiens.
 OS
 XX WO200160860-A2.
 PN
 XX 23-AUG-2001.
 XX
 XX 20-FEB-2001; 2001WO-US05171.
 XX
 PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 PR

XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 XX Schlegel R, Endege WO, Monahan JE;
 PI
 DR WPI; 2001-662795/76.
 XX
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer
 XX
 PS Claim 1; Page 1865; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX
 SQ Sequence 435 BP; 107 A; 100 C; 88 G; 140 T; 0 other;
 Query Match 56.0%; Score 14; DB 23; Length 435;
 Best Local Similarity 100.0%; Pred. NO. 2e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGTGCCAGGAACCA 14
 Db 218 TGTGCCAGGAACCA 205
 RESULT 44
 ABN75835/c
 ID ABN75835 standard; cDNA; 439 BP.
 XX
 AC ABN75835;
 XX
 DT 08-JUL-2002 (first entry)
 DE
 DE Human ORF782 cDNA, SEQ ID NO:1563.
 XX
 KW Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; haematopoiesis regulation; tissue growth;
 KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
 KW vasotropic; antipsoriatic; antidiabetic; cytostatic; neurotropic;
 KW neuroprotective; antithrombotic; antithrombotic; thrombolytic;
 KW cardiant; hypotensive; antithyroid; antinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200190366-A2.
 XX
 XX 29-NOV-2001.
 XX
 XX 24-MAY-2001; 2001WO-US17076.
 PF
 XX 24-MAY-2000; 2000US-206690P.
 PR

(CURA-) CURAGEN CORP.
 Leach MD, Shinkets RA;
 WPL: 2002-106200/14.
 P-PSDB; ABP31809.
 Novel human polypeptides and polynucleotides useful for diagnosing,
 preventing and treating cardiovascular disease, neurodegenerative,
 hyperproliferative disorders and disorders related to organ
 transplantation -
 Claim 1; Page 656; 2508pp; English.
 Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 ABN79587 represent cDNAs encoding them. The invention also encompasses
 polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 referred to as ORFX) proteins, polynucleotides at least 85% identical to
 the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 polynucleotides, the recombinant production of ORFX proteins, antibodies
 specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 polypeptides, methods of screening for modulators of ORFX expression or
 activity, and methods of screening individuals for a predisposition to an
 ORFX-associated disorder. The ORFX proteins of the invention have a wide
 range of biological activities, such as cytokine, cell proliferation,
 cell differentiation, immune modulation, haematopoiesis regulation,
 tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 chemokinetic activity, haemostatic activity, thrombolytic activity,
 receptor/ligand, antiinflammatory activity, tumour inhibition activity,
 and antifretive activity, and may also be involved in the determination
 of bodily characteristics, fertility and behaviour. ORFX proteins,
 nucleic acids and antibodies may be used in the treatment of cancers,
 other proliferative disorders such as psoriasis and benign tumours,
 neurological disorders such as epilepsy and Alzheimer's disease,
 cardiovascular diseases, immune system disorders, disorders related to
 organ transplantation, disorders of tissue growth and regeneration,
 diseases such as diabetes mellitus, hypothyroidism, and cholesterol
 storage disease, and infectious diseases caused by viral, bacterial,
 fungal and other pathogens. ORFX nucleic acids may also be used as a
 source of primers and probes, in the detection of ORFX genomic sequences
 or transcripts, in the identification and cloning of homologous
 sequences, in genetic diagnosis, and in forensic biology. The ORFX
 nucleic acids may additionally be used to produce transgenic animals
 which may be useful for studying the function and/or activity of ORFX
 protein, and in drug screening. The ORFX proteins may also be used as
 immunogens to generate specific antibodies, which are useful in the
 diagnosis, treatment and monitoring of ORFX-associated diseases.
 Sequence 439 BP; 114 A; 84 C; 112 G; 128 T; 1 other;
 Query Match 56.0%; Score 14; DB 24; Length 439;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 3 TGCCAGGAACACAGA 16
 |||||||
 Db 250 TGCCAGGAACACAGA 237
 RESULT 45
 AAK60019/c
 ID AAK60019 standard; cDNA; 454 BP.
 XX
 AC AAK60019;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:5079.
 XX
 DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis; ss.
 KW

XX Homo sapiens.
 OS WO200157182-A2.
 PN
 XX
 XX 09-AUG-2001.
 PD
 XX
 XX 17-JAN-2001; 2001WO-US01354.
 PF
 XX
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 14-JUL-2000; 2000US-0217496.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 18-AUG-2000; 2000US-0225759.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234597.
 PR 25-SEP-2000; 2000US-0234598.

PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0251989.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX

PA (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483426/52.
XX P-PSDB; AAM87238.
DR Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
PT Claim 1; SEQ ID NO 5079; 3071pp + Sequence Listing; English.
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent, the
CC cancers and cancer metastases of hematopoietic-related diseases, especially
CC to AAK87694 represent human immune/hematopoietic-antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 454 BP; 142 A; 70 C; 121 G; 118 T; 3 other;
Query Match 56.0%; Score 14; DB 22; Length 454;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GTGCCAGGACCCAG 15
Db 303 GTGCCAGGACCCAG 290
|||||

Search completed: November 24, 2002, 13:27:16
Job time : 855.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2002, 12:31:33 ; Search time 1611.5 Seconds
(without alignments)
451.486 Million cell updates/sec

Title: US-09-784-423-125

Perfect score: 25

Sequence: 1 TGTCACAGGACCAAGAAATTTACAG 25

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size: 0

Total number of hits satisfying chosen parameters: 4105280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg_hum.*

31: em.htg_inv.*

32: em.htg_other.*

33: em.htg_mus.*

34: em.htg_pin.*

35: em.htg_rod.*

36: em.htg_mam.*

37: em.htg_vrt.*

38: em.sy.*

39: em.htgo_hum.*

40: em.htgo_mus.*

41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	100.0	25	6	ARI54075
c	2	25	100.0	1000	ARI53982
c	3	25	100.0	62102	2
c	4	25	100.0	69964	9
c	5	25	100.0	138459	2
c	6	19	76.0	124427	9
c	7	19	76.0	157395	9
c	8	19	76.0	159767	2
c	9	19	76.0	227611	2
c	10	19	76.0	322058	2
c	11	18	72.0	133337	2
c	12	18	72.0	146310	9
c	13	18	72.0	169013	2
c	14	18	72.0	181380	10
c	15	18	72.0	182416	2
c	16	18	72.0	188787	2
c	17	18	72.0	191197	2
c	18	17	68.0	14333	8
c	19	17	68.0	40740	9
c	20	17	68.0	55252	2
c	21	17	68.0	104934	2
c	22	17	68.0	107361	2
c	23	17	68.0	110000	2
c	24	17	68.0	129151	9
c	25	17	68.0	145481	9
c	26	17	68.0	160538	2
c	27	17	68.0	174736	9
c	28	17	68.0	176861	9
c	29	17	68.0	185473	9
c	30	17	68.0	217873	9
c	31	17	68.0	220388	2
c	32	17	68.0	231923	2
c	33	17	68.0	257703	3
c	34	16	64.0	286	8
c	35	16	64.0	320	8
c	36	16	64.0	595	6
c	37	16	64.0	1147	8
c	38	16	64.0	1268	8
c	39	16	64.0	1358	8
c	40	16	64.0	1608	8
c	41	16	64.0	1700	4
c	42	16	64.0	4894	8
c	43	16	64.0	12558	8
c	44	16	64.0	15783	8
c	45	16	64.0	21941	9

ALIGNMENTS

RESULT 1	ARI54075	Sequence 125 from patent US 6238863.	25 bp	DNA	linear	PAT 08-AUG-2001
ARI54075	LOCUS					
DEFINITION	Sequence 125 from patent US 6238863.					
ACCESSION	ARI54075					
VERSION	ARI54075.1	GI:15122128				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 25)					
AUTHORS	Schumm, J.W. and Bacher, J.W.					
TITLE	Materials and methods for identifying and analyzing intermediate tandem repeat DNA markers					
JOURNAL	Patent: US 6238863-A 125 29-MAY-2001;					

```

FEATURES             Location/Qualifiers
     source            1..25
BASE COUNT           9 a 5 c 6 g 5 t
ORIGIN
Query Match          100.0%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0003;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGTGCCAGGAAACCAGAAATTTACAG 25
    |||
Db 1 TGTGCCAGGAAACCAGAAATTTACAG 25

RESULT 2
ARI53982/c
LOCUS               ARI53982
DEFINITION          Sequence 32 from patent US 6238863.
ACCESSION            ARI53982
VERSION              ARI53982.1 GI:15122035
KEYWORDS
SOURCE              Unknown.
ORGANISM            Unclassified.
REFERENCE            1 (bases 1 to 1000)
AUTHORS             Schumm,J.W. and Bacher,J.W.
TITLE              Materials and methods for indentifying and analyzing intermediate
                  tandem repeat DNA markers
JOURNAL             Patent: US 6238863-A 32 29-MAY-2001;
FEATURES             Location/Qualifiers
                  1..1000
source              /organism="unknown"
BASE COUNT          302 a 240 c 231 g 226 t 1 others
ORIGIN
Query Match          100.0%; Score 25; DB 6; Length 1000;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGTGCCAGGAAACCAGAAATTTACAG 25
    |||
Db 726 TGTGCCAGGAAACCAGAAATTTACAG 702

RESULT 3
AC104363/c
LOCUS               AC104363
DEFINITION          Homo sapiens chromosome 18 clone RP11-42N11 map 18, LOW-PASS
                  SEQUENCE SAMPLING.
ACCESSION            AC104363
VERSION              AC104363.1 GI:17426352
KEYWORDS             HTG; HTGS_PHASE0.
SOURCE              Homo sapiens.
ORGANISM            Homo sapiens
REFERENCE            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS             Birren,B., Linton,L., Nusbaum,C., Allen,N.,
                  Brown,A., Cantara,J., Campiano,A., Chang,J., Chazaro,B.,
                  Choepel,I., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
                  Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
                  Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
                  Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
                  Jones,C., Kanat,A., Karatas,A., Kells,C., LaRocque,K.,
                  Lamarcaes,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
                  MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,

```

McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,K., Rize,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strausman,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topnam,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE JOURNAL

COMMENT

Submitted (08-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WtBR
Web site: <http://www.seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L22094
Center Clone name: 42_N_11

* NOTE: This record contains 81 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 665: contig of 665 bp in length
666 765: gap of 100 bp
766 1454: contig of 689 bp in length
1455 1554: gap of 100 bp
1555 2232: contig of 678 bp in length
2233 2332: gap of 100 bp
2333 3028: contig of 696 bp in length
3029 3128: gap of 100 bp
3129 3753: contig of 625 bp in length
3754 3853: gap of 100 bp
3854 4459: contig of 606 bp in length
4460 4559: gap of 100 bp
4560 5225: contig of 666 bp in length
5226 5325: gap of 100 bp
5326 5979: contig of 654 bp in length
5980 6079: gap of 100 bp
6080 6752: contig of 673 bp in length
6753 6852: gap of 100 bp
6853 7526: contig of 674 bp in length
7527 7626: gap of 100 bp
7627 8307: contig of 681 bp in length
8308 8407: gap of 100 bp
8408 9090: contig of 683 bp in length
9091 9190: gap of 100 bp
9191 9880: contig of 690 bp in length
9881 9980: gap of 100 bp
9981 10667: contig of 687 bp in length
10668 10767: gap of 100 bp
10768 11436: contig of 669 bp in length
11437 11536: gap of 100 bp
11537 12157: contig of 621 bp in length
12158 12257: gap of 100 bp
12258 12913: contig of 656 bp in length
12914 13013: gap of 100 bp
13014 13667: contig of 654 bp in length
13668 13767: gap of 100 bp


```

* 13768 14429: contig of 662 bp in length
* 14430 14529: gap of 100 bp
* 14530 15189: contig of 660 bp in length
* 15190 15289: gap of 100 bp
* 15290 15950: contig of 661 bp in length
* 15951 16050: gap of 100 bp
* 16051 16724: contig of 674 bp in length
* 16725 16824: gap of 100 bp
* 16825 17505: contig of 681 bp in length
* 17506 17605: gap of 100 bp
* 17606 18290: contig of 685 bp in length
* 18291 18390: gap of 100 bp
* 18391 19050: contig of 660 bp in length
* 19051 19150: gap of 100 bp
* 19151 19807: contig of 657 bp in length
* 19808 19907: gap of 100 bp
* 19908 20568: contig of 661 bp in length
* 20569 20668: gap of 100 bp
* 20669 21322: contig of 654 bp in length
* 21323 21422: gap of 100 bp
* 21423 22079: contig of 657 bp in length
* 22080 22179: gap of 100 bp
* 22180 22852: contig of 673 bp in length
* 22853 22952: gap of 100 bp
* 22953 23621: contig of 669 bp in length
* 23622 23721: gap of 100 bp
* 23722 24386: contig of 665 bp in length
* 24387 24486: gap of 100 bp
* 24487 25174: contig of 688 bp in length
* 25175 25274: gap of 100 bp
* 25275 25957: contig of 683 bp in length
* 25958 26057: gap of 100 bp
* 26058 26729: contig of 672 bp in length
* 26730 26829: gap of 100 bp
* 26830 27478: contig of 649 bp in length
* 27479 27578: gap of 100 bp
* 27579 28237: contig of 659 bp in length
* 28238 28337: gap of 100 bp
* 28338 29002: contig of 665 bp in length
* 29003 29102: gap of 100 bp
* 29103 29764: contig of 662 bp in length
* 29765 29864: gap of 100 bp
* 29865 30544: contig of 680 bp in length
* 30545 30644: gap of 100 bp
* 30645 31330: contig of 686 bp in length
* 31331 31430: gap of 100 bp
* 31431 32107: contig of 677 bp in length
* 32108 32207: gap of 100 bp
* 32208 32887: contig of 680 bp in length
* 32888 32987: gap of 100 bp
* 32988 33679: contig of 692 bp in length
* 33680 33779: gap of 100 bp
* 33780 34451: contig of 672 bp in length
* 34452 34551: gap of 100 bp
* 34552 35216: contig of 665 bp in length
* 35217 35316: gap of 100 bp
* 35317 35978: contig of 662 bp in length
* 35979 36078: gap of 100 bp
* 36079 36731: contig of 653 bp in length
* 36732 36831: gap of 100 bp
* 36832 37486: contig of 655 bp in length
* 37487 37586: gap of 100 bp
* 37587 38249: contig of 663 bp in length
* 38250 38349: gap of 100 bp
* 38350 39032: contig of 683 bp in length
* 39033 39132: gap of 100 bp
* 39133 39815: contig of 683 bp in length
* 39816 39915: gap of 100 bp
* 39916 40576: contig of 661 bp in length
* 40577 40676: gap of 100 bp
* 40677 41356: contig of 680 bp in length
* 41357 41456: gap of 100 bp
* 41457 42141: contig of 665 bp in length

```

```

* 42142 42241: gap of 100 bp
* 42242 42899: contig of 658 bp in length
* 42900 42999: gap of 100 bp
* 43000 43650: contig of 651 bp in length
* 43651 43750: gap of 100 bp
* 43751 44412: contig of 662 bp in length
* 44413 44512: gap of 100 bp
* 44513 45169: contig of 657 bp in length
* 45170 45269: gap of 100 bp
* 45270 45925: contig of 656 bp in length
* 45926 46025: gap of 100 bp
* 46026 46695: contig of 670 bp in length
* 46696 46795: gap of 100 bp
* 46796 47481: contig of 686 bp in length
* 47482 47581: gap of 100 bp
* 47582 48185: contig of 604 bp in length
* 48186 48285: gap of 100 bp
* 48286 48964: contig of 679 bp in length
* 48965 49064: gap of 100 bp
* 49065 49751: contig of 687 bp in length
* 49752 49851: gap of 100 bp
* 49852 50518: contig of 667 bp in length
* 50519 50618: gap of 100 bp
* 50619 51290: contig of 672 bp in length
* 51291 51390: gap of 100 bp
* 51391 52048: contig of 658 bp in length
* 52049 52148: gap of 100 bp
* 52149 52826: contig of 678 bp in length

```

```

Query Match      100.0%; Score 25; DB 2; Length 62102;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TGTGCCAGCAGCAATTTTACAG 25
|||||

```

```

DB 55346 TGTGCCAGCAGCAATTTTACAG 55322
|||||

```

RESULT 4

HS496C20

LOCUS

DEFINITION

Human DNA sequence from clone RP3-496C20 on chromosome 22 Contains

STSs, GSSs and two putative CpG islands, complete sequence.

283847

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

HS496C20 69964 bp DNA linear PRI 12-DEC-1999
Human DNA sequence from clone RP3-496C20 on chromosome 22 Contains
STSs, GSSs and two putative CpG islands, complete sequence.

283847.1 GI:3550060

HTG; CpG island.

Homo sapiens.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 69964)

Clark G.

Direct Submission

Submitted (08-DEC-1999) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

requests: clonerequest@sanger.ac.uk

On Sep 8, 1998 this sequence version replaced gi:3413284.

as follows. An attempt is made to resolve all sequencing problems,

such as compressions and repeats, but not necessarily within known

annotated human repeat sequence elements (e.g. Alu). Where the

sequence is ambiguous, there is an annotation using the 'unsure'

feature key.

This sequence was generated from part of bacterial clone contigs of

human chromosome 22, constructed by the Sanger Centre Chromosome 22

Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr22>

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
 Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP3-496C20 is from the library RP3-3 constructed at the Roswell Park Cancer Institute by the group of Peter de Jong. For further details see <http://bacpac.med.buffalo.edu/>

VECTOR: pCYPAC2

This sequence is the entire insert of clone RP3-496C20. The true left end of clone RP3-340K22 is at 21972 in this sequence. The true right end of clone RP1-140N12 is at 41206 in this sequence. The start of this sequence overlaps with sequence 282206. The end of this sequence overlaps with sequence AL022238.

FEATURES

```

    Location/Qualifiers
      1..69964
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="22"
        /clone="RP3-496C20"
        /clone_lib="RPCI-3"
        24..186
      /note="match: STS: Em:HS496C20S; match: STS: Em:AL021673"
      169..311
        /note="MIR repeat: matches 71..211 of consensus"
      repeat_region
        551..600
      repeat_region
        /note="10 copies 5 mer ttttt 76 conserved"
      repeat_region
        2394..2462
      repeat_region
        /note="L2 repeat: matches 2626..2696 of consensus"
      repeat_region
        3060..3221
      repeat_region
        /note="MIR repeat: matches 21..192 of consensus"
      repeat_region
        3226..3535
      repeat_region
        /note="AluJb repeat: matches 5..312 of consensus"
      repeat_region
        4446..4662
      repeat_region
        /note="MIR repeat: matches 35..236 of consensus"
      repeat_region
        5924..5970
      repeat_region
        /note="L2 repeat: matches 2701..2748 of consensus"
      misc_feature
        6015..6330
      repeat_region
        /note="match: GSS: Em:AQ356083"
      repeat_region
        6149..6461
      repeat_region
        /note="AluX repeat: matches 1..312 of consensus"
      repeat_region
        6462..6596
      repeat_region
        /note="AluSq/x repeat: matches 1..135 of consensus"
      repeat_region
        6632..6695
      repeat_region
        /note="MIR repeat: matches 70..146 of consensus"
      repeat_region
        7617..8113
      repeat_region
        /note="AluX repeat: matches 1..297 of consensus"
      misc_feature
        complement(8207..8891)
      repeat_region
        /note="match: GSS: Em:AQ344529"
      repeat_region
        8840..8926
      repeat_region
        /note="MIR repeat: matches 80..191 of consensus"
      repeat_region
        9079..9606
      repeat_region
        /note="264 copies 2 mer cc 58 conserved"
      misc_feature
        9187..9189
      repeat_region
        /note="forced join; gap estimated as 1kb by restriction digest data"
      misc_feature
        complement(9187..9189)
      repeat_region
        /note="Forced join. Gap estimated as 1kb by restriction digest data."
      repeat_region
        9481..9605
      repeat_region
        /note="25 copies 5 mer ctctt 62 conserved"
      repeat_region
        10240..10395
      repeat_region
        /note="MIR repeat: matches 76..255 of consensus"
      repeat_region
        10502..10587
      repeat_region
        /note="MER5A repeat: matches 57..150 of consensus"
      repeat_region
        10511..10643
      repeat_region
        /note="MER5A repeat: matches 57..185 of consensus"
      repeat_region
        11303..11539
      repeat_region
        /note="L2 repeat: matches 1865..2127 of consensus"
      repeat_region
        13046..13314
      repeat_region
        /note="AluX repeat: matches 28..305 of consensus"
      repeat_region
        13527..13983
      repeat_region
        /note="L1M8 repeat: matches 5393..5854 of consensus"
      repeat_region
        13984..14293
      repeat_region
        /note="AluSq repeat: matches 1..310 of consensus"
      repeat_region
        14294..14347
      repeat_region
        /note="L1M8 repeat: matches 5340..5393 of consensus"
      repeat_region
        14348..14649
      repeat_region
        /note="AluX repeat: matches 2..304 of consensus"
      repeat_region
        14651..14966
      repeat_region
        /note="AluY repeat: matches 1..306 of consensus"
      repeat_region
        14967..15007
      repeat_region
        /note="L1M8 repeat: matches 5352..5341 of consensus"
      repeat_region
        15024..16371
      repeat_region
        /note="L1M4 repeat: matches 3784..5097 of consensus"
      repeat_region
        16413..16611
      repeat_region
        /note="L1 repeat: matches 3186..3382 of consensus"
      repeat_region
        16662..16875
      repeat_region
        /note="MIR repeat: matches 29..256 of consensus"
      repeat_region
        16891..17310
      repeat_region
        /note="MLT2FB repeat: matches 2..414 of consensus"
      misc_feature
        17137..17686
      repeat_region
        /note="match: GSS: Em:AQ551353"
      repeat_region
        17359..17438
      repeat_region
        /note="MER91A repeat: matches 114..193 of consensus"
      repeat_region
        17825..18044
      repeat_region
        /note="LTR23 repeat: matches 210..437 of consensus"
      repeat_region
        18046..18357
      repeat_region
        /note="AluSq repeat: matches 1..309 of consensus"
      repeat_region
        18511..18822
      repeat_region
        /note="AluSq repeat: matches 1..313 of consensus"
      repeat_region
        18830..18897
      repeat_region
        /note="BURL repeat: matches 6300..6367 of consensus"
      repeat_region
        19044..19080
      repeat_region
        /note="MIR repeat: matches 207..245 of consensus"
      repeat_region
        19081..19391
      repeat_region
        /note="AluSq repeat: matches 1..307 of consensus"
      repeat_region
        19392..19492
      repeat_region
        /note="MIR repeat: matches 117..207 of consensus"
      repeat_region
        19581..19801
      repeat_region
        /note="MIR repeat: matches 29..261 of consensus"
      repeat_region
        20302..20464
      repeat_region
        /note="AluSq/x repeat: matches 134..302 of consensus"
      repeat_region
        20465..20747
      repeat_region
        /note="AluSq repeat: matches 1..296 of consensus"
      repeat_region
        22217..22508
      repeat_region
        /note="AluX repeat: matches 14..305 of consensus"
      repeat_region
        22764..22848
      repeat_region
        /note="MIR repeat: matches 164..249 of consensus"
      misc_feature
        23209..23725
      repeat_region
        /note="CpG island"
      repeat_region
        /evidence=not_experimental
      repeat_region
        23946..24063
      repeat_region
        /note="MIR repeat: matches 21..140 of consensus"
      repeat_region
        25053..25142
      repeat_region
        /note="MIR repeat: matches 21..119 of consensus"
      repeat_region
        25207..25233
      repeat_region
        /note="9 copies 3 mer gct 96 conserved"
      misc_feature
        25401..26038
      repeat_region
        /note="match: GSS: Em:B54652"
      repeat_region
        25877..26155
      repeat_region
        /note="AluJb repeat: matches 1..296 of consensus"
      repeat_region
        26997..27304
      repeat_region
        /note="AluJb repeat: matches 1..311 of consensus"
      repeat_region
        27326..27595
      repeat_region
        /note="AluX repeat: matches 37..311 of consensus"
      repeat_region
        28098..28406
      repeat_region
        /note="AluSq repeat: matches 1..311 of consensus"
      repeat_region
        28847..29128
      repeat_region
        /note="L2 repeat: matches 2362..2745 of consensus"
      misc_feature
        complement(29031..29545)
      repeat_region
        /note="match: GSS: Em:AQ528941"
      repeat_region
        30940..30989
      repeat_region
        /note="25 copies 2 mer gt 100 conserved"
      repeat_region
        30990..31054
      repeat_region
        /note="MER41A repeat: matches 400..464 of consensus"

```

```

misc_feature      31120..31339
                    /note="match: STS: Em:G03731; match: STS: Em:G03731"
repeat_region     34416..34721
                    /note="AluY repeat: matches 1..310 of consensus"
repeat_region     34776..34821
                    /note="L2 repeat: matches 2688..2733 of consensus"
repeat_region     34944..35243
                    /note="AluX repeat: matches 1..300 of consensus"
repeat_region     35351..35551
                    /note="MER20 repeat: matches 1..204 of consensus"
repeat_region     36460..36544
                    /note="L2 repeat: matches 2665..2749 of consensus"
repeat_region     37480..37620
                    /note="MIR repeat: matches 4..150 of consensus"
repeat_region     37861..38188
                    /note="AluJo repeat: matches 1..312 of consensus"
                    complement(38024..38419)
misc_feature       /note="match: STS: Em:G28177; match: STS: Em:R54672"
misc_feature       /note="match: STS: Em:G25989"
repeat_region     38802..38948

```

```

Query Match      100.0%; Score 25; DB 9; Length 69964;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCGAGACACGAAATTTACAG 25
    |||||
Db 5959 TGTGCGAGACACGAAATTTACAG 5983

```

```

RESULT 5
AC090795/C
LOCUS              AC090795              138459 bp    DNA        linear   HTG 11-JUN-2001
DEFINITION         Homo sapiens chromosome 8 clone RP11-254K5 map 8, WORKING DRAFT
SEQUENCE           AC090795              12 unordered pieces.
ACCESSION          AC090795
VERSION            AC090795.2   GT:14336499
KEYWORDS           HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE             Homo sapiens
ORGANISM           Homo sapiens
REFERENCE
AUTHORS            Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
                    Barua,N., Bastien,V., Boguslavsky,L., Boukhgalter,B., Brown,A.,
                    Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
                    Collymore,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S.,
                    Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
                    Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
                    Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
                    Jones,C., Karatas,A., LaRoque,K., Lamazares,R., Landers,T.,
                    Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
                    Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
                    McPheeters,R., Meldrum,J., Meneus,L., Mihova,I., Mlenga,V.,
                    Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
                    O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
                    Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
                    Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
                    Roy,A., Santos,R., Schauer,S., Schuppback,R., Seaman,S., Severy,P.,
                    Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
                    Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
                    Travers,M., Travis,N., Triggilio,J., Vassiliev,H., Viel,R., Vo,A.,
                    Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
                    Zembek,L., Zimmer,A. and Zody,M.
DIRECT SUBMISSION
TITLE              Submitted (10-MAR-2001) Whitehead Institute/MIT Center for Genome
JOURNAL            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT            On Jun 11, 2001 this sequence version replaced gi:13270664.

```

All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L12790
 Center clone name: 254_K 5
 ----- Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 133567 bases at least Q40
 Consensus quality: 135811 bases at least Q30
 Consensus quality: 136678 bases at least Q20
 Insert size: 139000; agarose-fp
 Quality coverage: 6.5 in Q20 bases; agarose-fp
 Quality coverage: 6.6 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 12 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

* 1 1411: contig of 1411 bp in length
* 1412 1511: gap of 100 bp
* 1512 2165: contig of 654 bp in length
* 2166 2265: gap of 100 bp
* 2266 3074: contig of 809 bp in length
* 3075 3174: gap of 100 bp
* 3175 4763: contig of 1589 bp in length
* 4764 4863: gap of 100 bp
* 4864 6161: contig of 1298 bp in length
* 6162 6261: gap of 100 bp
* 6262 8601: contig of 2340 bp in length
* 8602 8701: gap of 100 bp
* 8702 13139: contig of 4438 bp in length
* 13140 13239: gap of 100 bp
* 13240 17613: contig of 4374 bp in length
* 17614 17713: gap of 100 bp
* 17714 30894: contig of 13181 bp in length
* 30895 30994: gap of 100 bp
* 30995 40051: contig of 9057 bp in length
* 40052 40151: gap of 100 bp
* 40152 95859: contig of 55708 bp in length
* 95860 95959: gap of 100 bp
* 95960 138459: contig of 42500 bp in length.
FEATURES
    source
        1..138459
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="8"
            /map="8"
            /clone="RP11-254K5"
            /clone_lib="RPC1-11 Human Male BAC"
    misc_feature
        1..1411
            /note="assembly_fragment"
            clone_end:Sp6
            vector_side:left
        1512..2165
            /note="assembly_fragment"
        2266..3074
            /note="assembly_fragment"
        3175..4763
            /note="assembly_fragment"
        4864..6161
            /note="assembly_fragment"

```

```

misc_feature      /note="assembly_fragment"
6262..8601
/note="assembly_fragment"
8702..13139
/note="assembly_fragment"
13240..17613
/note="assembly_fragment"
17714..30894
/note="assembly_fragment"
30995..40051
/note="assembly_fragment"
40152..95859
/note="assembly_fragment"
95960..138459
/note="assembly_fragment
clone_end:77
vector_side:right"
BASE COUNT      38140 a 31260 c 30041 g 37914 t 1104 others
ORIGIN

Query Match      100.0%; Score 25; DB 2; Length 138459;
Best Local Similarity 100.0%; Pred. NO. 0.00016;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TGTGCCAGGACACGAAATTTACAG 25
|||||
Db 34265 TGTGCCAGGACACGAAATTTACAG 34241

RESULT 6
AC048344/c
LOCUS
DEFINITION      Homo sapiens 12 BAC RP11-843B15 (Roswell park Cancer Institute
ACCESSION      AC048344
VERSION
KEYWORDS
SOURCE
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Ate,J.R., Ayale,M., Banks,T.,
Barbata,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,F.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Devilla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S.,
Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P.,
Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N.,
Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Haviak,P.,
Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M.,
Holloway,C., Hollins,B., Homsli,F., Howard,S., Huber,J., Hulyk,S.,
Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y.,
Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,J.,
King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N.,
Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O.,
Lieu,C., Liu,J., Liu,W., Louisied,H., Lozado,R.J., Lu,X.,
Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P.,
Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E.,
Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S.,
Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T.,
Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M.,
Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Owiedo,K., Pace,A., Payton,B., Peery,J., Perez,L.,

```

Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C., Shooshtari,N., Sisson,I., Sodergren,E., Sonaik,E., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wang,Q., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R., Weinstock,G. and Gibbs,R.

Direct Submission

Unpublished
2 (bases 1 to 124427)
Worley,K.C.

Direct Submission

Submitted (14-APR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 124427)
Worley,K.C.

Direct Submission

Submitted (17-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 124427)
Worley,K.C.

Direct Submission

Submitted (25-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 124427)
Worley,K.C.

Direct Submission

Submitted (27-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jun 25, 2002 this sequence version replaced gi:21426084.

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annot>

```

FEATURES          ation.html.
source
  1. 124427
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /chromosome="12"
  /clone="rp11-843B15"
  127. 223
  /rpt_family="AluY"
  148. 488
  /standard_name="L1M3"
  /db_xref="GDB:631802"
  complement(226. 509)
  /rpt_family="AluSx"
  650. 776
  repeat_region
  /rpt_family="L1M3"
  825. 1055
  repeat_region
  /rpt_family="L1M3"
  1057. 1339
  repeat_region
  /rpt_family="AluSx"
  1340. 1644
  repeat_region
  /rpt_family="AluSp"
  1369. 1503
  STS
  /standard_name="46230"
  1645. 1779
  repeat_region
  /rpt_family="L1MA9"
  1825. 1939
  repeat_region
  /rpt_family="FLAM_C"
  1940. 1992
  repeat_region
  /rpt_family="TTTA"
  complement(1993. 2272)
  repeat_region
  /rpt_family="AluY"
  complement(2276. 2572)
  repeat_region
  /rpt_family="AluSx"
  2576. 2602
  repeat_region
  /rpt_family="AT-rich"
  2689. 2806
  repeat_region
  /rpt_family="FLAM_C"
  2817. 2838
  repeat_region
  /rpt_family="AT-rich"
  2839. 3137
  repeat_region
  /rpt_family="AluSc"
  3817. 3922
  repeat_region
  /rpt_family="MER113"
  complement(3950. 4047)
  repeat_region
  /rpt_family="MIR"
  4048. 4331
  repeat_region
  /rpt_family="AluSx"
  complement(4398. 4458)
  repeat_region
  /rpt_family="MIR"
  complement(4541. 4839)
  repeat_region
  /rpt_family="AluY"
  4912. 5306
  repeat_region
  /rpt_family="MER4A2"
  5307. 5587
  repeat_region
  /rpt_family="AluSx"
  5588. 5666
  repeat_region
  /rpt_family="MER4A2"
  5799. 6087
  repeat_region
  /rpt_family="AluSx"
  complement(6100. 6180)
  repeat_region
  /rpt_family="MER34"
  complement(6241. 6419)
  repeat_region
  /rpt_family="MER34B"
  6489. 6670
  repeat_region
  /rpt_family="CT-rich"
  complement(6674. 6755)
  repeat_region
  /rpt_family="MER34B"
  7008. 7139
  repeat_region
  /rpt_family="FLAM_C"
  7147. 7449
  repeat_region
  /rpt_family="AluSx"
  7450. 7494

```

```

/rpt_family="(CA)n"
7585. 8030
/rpt_family="L1MB8"
8049. 8346
/rpt_family="AluY"
8355. 8626
/rpt_family="AluJo"
complement(8678. 8806)
/rpt_family="MER110"
8855. 9151
/rpt_family="AluY"
complement(9179. 9328)

Query Match      76.0%; Score 19; DB 9; Length 124427;
Best Local Similarity 100.0%; Pred. No. 0.65;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AGGAACCAAGAAATTTACAG 25
|||||
Db 101453 AGGAACCAAGAAATTTACAG 101435

RESULT 7
CNS07EF8/c      157395 bp      DNA      linear      PRI 04-MAY-2001
LOCUS           Human chromosome 14 DNA sequence BAC R-286B5 of library RPT-11
DEFINITION      from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION       AL512359
VERSION         AL512359.2 GI:13992194
KEYWORDS        HTG.
SOURCE          human.
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Heilig, R., Petit, J.L., Vico, V., Dasilva, C., Robert, C., Wincker, P.,
Brotier, P., Cattolico, L., Barbe, V., Pelletier, E., Artiguenave, F.,
Levy, M., Eckenberg, R., Bruls, T., Deserardis, V., Cruaud, C.,
Sygpay, G., Saurin, W. and Weissbach, J.
Sequencing of the human chromosome 14
Unpublished
2 (bases 1 to 157395)
Genoscope.
Direct Submission
Submitted (04-MAY-2001) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
On May 8, 2001 this sequence version replaced gi:11990874.
----- Genoscope Center
Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: Seqref@genoscope.cns.fr
-----
The following BAC sequence is oriented from the T7 to the SP6 end.
Upstream BAC (overlapping the T7 end) : R-2G1
Downstream BAC (overlapping the SP6 end) : R-1109N18_PCR1
(AC-AL583762) ----- Summary Statistics
Assembly program: Phrap; version 2.0
Quality coverage: 6.59x in Q20 bases; sum-of-contigs
-----
Overall quality chart :
Range : bases
0 : 11
1 - 9 : 10
10 - 19 : 145
20 - 29 : 279
30 - 39 : 660
40 - 49 : 4075
50 - 59 : 7108
60 - 69 : 11941
70 - 79 : 27134

```

```

80 - 89 : 50220
90 - 99 : 55812
-----
Percentage of bases with a quality value >= 40 : 99 %.
FEATURES
  source
    Location/Qualifiers
      1. .157395
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="14"
        /clone="R-286B5"
        /clone_lib="RPCI-11"
BASE COUNT  52242 a 27541 c 27367 g 50245 t
ORIGIN
Query Match      76.0%; Score 19; DB 9; Length 157395;
Best Local Similarity 100.0%; Pred. No. 0.63;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY  1 TGTGCCAGAACCCAGAAAT 19
      |||||
Db  46581 TGTGCCAGAACCCAGAAAT 46563

RESULT 8
AC023529
LOCUS      Homo sapiens clone RP11-27P2, WORKING DRAFT SEQUENCE, 14 unordered
DEFINITION
ACCESSION  AC023529
VERSION    AC023529.3 GI:10047793
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE     Homo sapiens.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 159767)
AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE     Homo sapiens, clone RP11-27P2
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 159767)
AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavskiy,L.,
            Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
            Dearellano,K., Dewar,K., Dodge,S., Domino,M., Doyle,M.,
            Fenesor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., Landers,T., Largocque,K., Lehoczy,J., Levine,R.,
            Lieu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N., McCarthy,M.,
            McEwan,P., McGurk,A., McKernan,K., McPheters,R., Meldrum,J.,
            Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Olivier,T.M.,
            Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
            Riley,K., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
            Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
            Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.,
            Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
            Zody,M.
            Direct Submission
TITLE     Submitted (15-FEB-2000) Whitehead Institute/MIT Center for Genome
JOURNAL   Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT   On Sep 10, 2000 this sequence version replaced gi:7209947.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ----- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu

```

```

----- Project Information
Center project name: L4732
Center clone name: 27_P_2
----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Assembly: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 147596 bases at least Q40
Consensus quality: 154530 bases at least Q30
Consensus quality: 156916 bases at least Q20
Insert size: 164000; agarose-fp
Insert size: 158467; sum-of-contigs
Quality coverage: 3.7 in Q20 bases; agarose-fp
Quality coverage: 3.9 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 38585: contig of 38585 bp in length
* 38586 38685: gap of 100 bp
* 38686 40414: contig of 1729 bp in length
* 40415 40514: gap of 100 bp
* 40515 44702: contig of 4188 bp in length
* 44703 44802: gap of 100 bp
* 44803 49260: contig of 4458 bp in length
* 49261 49360: gap of 100 bp
* 49361 54772: contig of 5412 bp in length
* 54773 54872: gap of 100 bp
* 54873 61094: contig of 6222 bp in length
* 61095 61194: gap of 100 bp
* 61195 69615: contig of 8421 bp in length
* 69616 69715: gap of 100 bp
* 69716 76018: contig of 6303 bp in length
* 76019 76118: gap of 100 bp
* 76119 84944: contig of 8826 bp in length
* 84945 85044: gap of 100 bp
* 85045 94808: contig of 9764 bp in length
* 94809 94908: gap of 100 bp
* 94909 105344: contig of 10436 bp in length
* 105345 105444: gap of 100 bp
* 105445 117738: contig of 12294 bp in length
* 117739 117838: gap of 100 bp
* 117839 153494: contig of 35656 bp in length
* 153495 153594: gap of 100 bp
* 153595 159767: contig of 6173 bp in length.
FEATURES
  source
    Location/Qualifiers
      1. .159767
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="RP11-27P2"
        /clone_lib="RPCI-11 Human Male BAC"
        1. .38585
          /note="assembly_fragment"
          clone_end:SP6
          vector_side:left
          38686..40414
            /note="assembly_fragment"
            40515..44702
              /note="assembly_fragment"
              44803..49260
                /note="assembly_fragment"
                49361..54772
                  /note="assembly_fragment"
                  54873..61094
                    /note="assembly_fragment"
                    61195..69615
                      /note="assembly_fragment"
                      69716..76018

```

```

/misc_feature      /note="assembly_fragment"
76119. .84944
/misc_feature      /note="assembly_fragment"
85045. .94808
/misc_feature      /note="assembly_fragment"
94909. .105344
/misc_feature      /note="assembly_fragment"
105445. .117736
/misc_feature      /note="assembly_fragment"
117839. .153494
/misc_feature      /note="assembly_fragment"
153595. .159767
/misc_feature      /note="assembly_fragment"
clone_end:T7
vector_side:right"
BASE COUNT      50843 a 28235 c 28248 g 51140 t 1301 others
ORIGIN
Query Match      76.0%; Score 19; DB 2; Length 159767;
Best Local Similarity 100.0%; Pred. No. 0.63;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGTGCCAGGACACAGAAAT 19
|||||
Db 38092 TGTGCCAGGACACAGAAAT 38110

RESULT 9
AC023766/c
LOCUS      AC023766      227611 bp      DNA      linear      HTG 24-AUG-2002
DEFINITION Homo sapiens chromosome 12 clone RP11-459A22 map 12, WORKING DRAFT
SEQUENCE      AC023766      62 unordered pieces.
ACCESSION      AC023766
VERSION      AC023766.2      GI:7767814
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 227611)
AUTHORS      Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE      Homo sapiens chromosome 12, clone RP11-459A22
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 227611)
AUTHORS      Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavskiy,L.,
Bouckgaalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
Dearellano,K., Dewar,K., Dodge,S., Domino,M., Doyle,M.,
Fenster,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Landers,T., Largocque,K., Lehoczy,J., Levine,R.,
Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.M.,
Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.,
Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
Zody,M.
Direct Submission
Submitted (17-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 227611)
REFERENCE      3 (bases 1 to 227611)
AUTHORS      Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavskiy,L., Bouckgaalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,

```

Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Peterson,T.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE JOURNAL

COMMENT

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 10, 2000 this sequence version replaced gi:6984473.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L6351

Center clone name: 459_A.22

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 161281 bases at least Q40

Consensus quality: 183874 bases at least Q30

Consensus quality: 202702 bases at least Q20

Insert size: 162000; agarose-fp

Insert size: 221511; sum-of-contigs

Quality coverage: 4.9 in Q20 bases; agarose-fp

Quality coverage: 3.6 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 62 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

1 1178: contig of 1178 bp in length

1179 1278: gap of 100 bp

1279 2325: contig of 1047 bp in length

2326 2425: gap of 100 bp

2426 3504: contig of 1079 bp in length

3505 3604: gap of 100 bp

3605 4812: contig of 1208 bp in length

4813 4912: gap of 100 bp

4913 6336: contig of 1424 bp in length

6337 6436: gap of 100 bp

6437 7606: contig of 1170 bp in length

7607 7706: gap of 100 bp

7707 9220: contig of 1514 bp in length

9221 9320: gap of 100 bp

9321 10404: contig of 1084 bp in length

10405 10504: gap of 100 bp

10505 11787: contig of 1283 bp in length

11788 11887: gap of 100 bp

11888 13396: contig of 1509 bp in length

13397 13496: gap of 100 bp

13497 14524: contig of 1028 bp in length

* 14525 14624: gap of 100 bp
 * 14625 15925: contig of 1301 bp in length
 * 15926 16025: gap of 100 bp
 * 16026 16427: contig of 402 bp in length
 * 16428 16527: gap of 100 bp
 * 16528 17984: contig of 1457 bp in length
 * 17985 18084: gap of 100 bp
 * 18085 19857: contig of 1773 bp in length
 * 19858 19957: gap of 100 bp
 * 19958 21461: contig of 1504 bp in length
 * 21462 21561: gap of 100 bp
 * 21562 23157: contig of 1596 bp in length
 * 23158 23257: gap of 100 bp
 * 23258 24526: contig of 1269 bp in length
 * 24527 24626: gap of 100 bp
 * 24627 25955: contig of 1329 bp in length
 * 25956 26055: gap of 100 bp
 * 26056 27976: contig of 1921 bp in length
 * 27977 28076: gap of 100 bp
 * 28077 29874: contig of 1798 bp in length
 * 29875 29974: gap of 100 bp
 * 29975 31504: contig of 1530 bp in length
 * 31505 31604: gap of 100 bp
 * 31605 32976: contig of 1372 bp in length
 * 32977 33076: gap of 100 bp
 * 33077 34237: contig of 1161 bp in length
 * 34238 34337: gap of 100 bp
 * 34338 35654: contig of 1317 bp in length
 * 35655 35754: gap of 100 bp
 * 35755 37631: contig of 1877 bp in length
 * 37632 37731: gap of 100 bp
 * 37732 39294: contig of 1563 bp in length
 * 39295 39394: gap of 100 bp
 * 39395 41364: contig of 1970 bp in length
 * 41365 41464: gap of 100 bp
 * 41465 43336: contig of 1872 bp in length
 * 43337 43436: gap of 100 bp
 * 43437 45488: contig of 2052 bp in length
 * 45489 45588: gap of 100 bp
 * 45589 47902: contig of 2314 bp in length
 * 47903 48002: gap of 100 bp
 * 48003 50923: contig of 2921 bp in length
 * 50924 51023: gap of 100 bp
 * 51024 53227: contig of 2204 bp in length
 * 53228 53327: gap of 100 bp
 * 53328 56721: contig of 3394 bp in length
 * 56722 56821: gap of 100 bp
 * 56822 58718: contig of 1857 bp in length
 * 58719 58818: gap of 100 bp
 * 58819 61916: contig of 3098 bp in length
 * 61917 62016: gap of 100 bp
 * 62017 64899: contig of 2883 bp in length
 * 64900 64999: gap of 100 bp
 * 65000 69476: contig of 4477 bp in length
 * 69477 69576: gap of 100 bp
 * 69577 74433: contig of 4857 bp in length
 * 74434 74533: gap of 100 bp
 * 74534 78765: contig of 4232 bp in length
 * 78766 78865: gap of 100 bp
 * 78866 83783: contig of 4918 bp in length
 * 83784 83883: gap of 100 bp
 * 83884 88554: contig of 4671 bp in length
 * 88555 88654: gap of 100 bp
 * 88655 92887: contig of 4233 bp in length
 * 92888 92987: gap of 100 bp
 * 92988 97174: contig of 4187 bp in length
 * 97175 97274: gap of 100 bp
 * 97275 101602: contig of 4328 bp in length
 * 101603 101702: gap of 100 bp
 * 101703 107534: contig of 5832 bp in length
 * 107535 107634: gap of 100 bp
 * 107635 112107: contig of 4473 bp in length
 * 112108 112207: gap of 100 bp

* 112208 117172: contig of 4965 bp in length
 * 117173 117272: gap of 100 bp
 * 117273 121060: contig of 3788 bp in length
 * 121061 121160: gap of 100 bp
 * 121161 124846: contig of 3686 bp in length
 * 124847 124946: gap of 100 bp
 * 124947 129670: contig of 4724 bp in length
 * 129671 129770: gap of 100 bp
 * 129771 134148: contig of 4378 bp in length
 * 134149 134248: gap of 100 bp

Query Match 76.0%; Score 19; DB 2; Length 227611;
 Best Local Similarity 100.0%; Pred. No. 0.62;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AGGAACACAGAAATTACAG 25
 Db 194040 AGGAACACAGAAATTACAG 194022

RESULT 10

AC068887

LOCUS

DEFINITION

Homo sapiens chromosome 12 clone RP11-392G11, WORKING DRAFT

SEQUENCE, 12 unordered pieces.

AC068887

AC068887

HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP; HTGS_ACTIVEFIN.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 322058)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-ouman,F.R., Allen,C.,
 Alsbrooks,S.L., Amarantunga,H.C., Are,J.R., Ayele,M., Banks,T.,
 Barbara,J., Benton,J., Blincoe,K., Blankenburg,K., Bonnin,D.,
 Bouck,J., Bowie,S., Briviera,M., Brown,E., Brown,M., Bryant,N.P.,
 Buhay,C., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 DeLunay,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falls,T., Ferraguto,B., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabisi,A., Geo,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
 Hernandez,O., Hodgson,A., Hoques,M., Holloway,C., Hollins,B.,
 Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudan,S.,
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulsegad,H.,
 Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
 Massey,E., Mathewney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morigan,M., Morris,S.,
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
 Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okwuonu,G.,
 Oragunye,N., Owiedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
 Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
 Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,
 Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H.,
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
 Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
 Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstein,G. and Gibbs,R.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Direct Submission
Unpublished
2 (bases 1 to 322058)
Worley, K.C.
Direct Submission
Submitted (11-MAY-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 322058)
Worley, K.C.
Direct Submission
Submitted (29-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jun 25, 2002 this sequence version replaced gi:21535845.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Drafting Center code: BCM
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HAQU
Center clone name: RP11-392G11
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-primer Bodypy; 1% of reads
Chemistry: Dye-terminator Big Dye; 99% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 314551 bases at least Q40
Consensus quality: 316234 bases at least Q30
Consensus quality: 316974 bases at least Q20
Estimated insert size: 309838; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

COMMENT

***** NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
***** NOTE: This sequence may represent more than one clone.
***** NOTE: This is a 'working draft' sequence. It currently
consists of 12 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
***** This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 2423: contig of 2423 bp in length
2424 2523: gap of unknown length
2524 8179: contig of 5656 bp in length
8180 8297: gap of unknown length
8298 22503: contig of 14224 bp in length
22504 22603: gap of unknown length
22604 35216: contig of 12613 bp in length
35217 35316: gap of unknown length
35317 42339: contig of 7023 bp in length
42340 42439: gap of unknown length
42440 62720: contig of 20281 bp in length
62721 62820: gap of unknown length
62821 71459: contig of 8639 bp in length
71460 71559: gap of unknown length
71560 109808: contig of 38249 bp in length
109809 109908: gap of unknown length
109909 134937: contig of 25029 bp in length
134938 135037: gap of unknown length
135038 164861: contig of 29824 bp in length
164862 164961: gap of unknown length
164962 235517: contig of 70556 bp in length
235518 235617: gap of unknown length
235618 322058: contig of 86441 bp in length.
Location/Qualifiers
1. 322058
/organism="Homo sapiens"

FEATURES
Source

/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-392G11"
BASE COUNT 98314 a 65757 c 63670 g 93215 t 1102 others
ORIGIN
Query Match 76.0%; Score 19; DB 2; Length 322058;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 AGGAACCCAGAAATTACAG 25
|||||
Db 248175 AGGAACCCAGAAATTACAG 248193
|||||
RESULT 11
AC016231/c
LOCUS
DEFINITION Homo sapiens clone RP11-27H13, WORKING DRAFT SEQUENCE, 39 unordered
pieces.
ACCESSION AC016231
VERSION AC016231.4 GI:13123345
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 133337)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens, clone RP11-27H13
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 133337)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Barna, N., Beckerly, R., Boquslavsky, L., Boukhalter, B.,
Brown, A., Castle, A., Colangelo, M., Collins, S., Collamore, A.,
Cooke, P., DeAtellano, K., Dewar, K., Domino, M., Donegan, L., Doyle, M.,
Ferreira, P., FitzHugh, W., Forrest, C., Funk, R., Gage, D.,
Galaan, J., Gardyna, S., Grant, G., Hago, B., Headford, A., Horton, L.,
Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Lenczky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,
McEwan, P., McGurk, A., McKernan, K., McDonald, P., Marquis, N.,
Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (23-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 25, 2001 this sequence version replaced gi:12039543.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L4697
Center clone name: 27_H13
----- Summary Statistics
Sequencing vector: M13; M77815; 11% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 116391 bases at least Q40
Consensus quality: 125269 bases at least Q30
Consensus quality: 128158 bases at least Q20
Insert size: 153000; agarose-fp
Quality coverage: 2.6 in Q20 bases; agarose-fp
Quality coverage: 3.1 in Q20 ba.

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 39 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1 672: contig of 672 bp in length
* 673 772: gap of 100 bp
* 773 1578: contig of 806 bp in length
* 1579 1678: gap of 100 bp
* 1679 13396: contig of 11718 bp in length
* 13397 13496: gap of 100 bp
* 13497 14553: contig of 1057 bp in length
* 14554 14653: gap of 100 bp
* 14654 15808: contig of 1155 bp in length
* 15809 15908: gap of 100 bp
* 15909 17297: contig of 1389 bp in length
* 17298 17397: gap of 100 bp
* 17398 18345: contig of 948 bp in length
* 18346 18445: gap of 100 bp
* 18446 19764: contig of 1319 bp in length
* 19765 19864: gap of 100 bp
* 19865 21198: contig of 1334 bp in length
* 21199 21298: gap of 100 bp
* 21299 22590: contig of 1292 bp in length
* 22591 22690: gap of 100 bp
* 22691 24188: contig of 1498 bp in length
* 24189 24288: gap of 100 bp
* 24289 25165: contig of 877 bp in length
* 25166 25265: gap of 100 bp
* 25266 27286: contig of 2021 bp in length
* 27287 27386: gap of 100 bp
* 27387 29216: contig of 1830 bp in length
* 29217 29316: gap of 100 bp
* 29317 32525: contig of 3209 bp in length
* 32526 32625: gap of 100 bp
* 32626 34252: contig of 1627 bp in length
* 34253 34352: gap of 100 bp
* 34353 37416: contig of 3064 bp in length
* 37417 37516: gap of 100 bp
* 37517 40514: contig of 2998 bp in length
* 40515 40614: gap of 100 bp
* 40615 44387: contig of 3773 bp in length
* 44388 44487: gap of 100 bp
* 44488 46779: contig of 2292 bp in length
* 46780 46879: gap of 100 bp
* 46880 49231: contig of 2352 bp in length
* 49232 49331: gap of 100 bp
* 49332 52593: contig of 3262 bp in length
* 52594 52693: gap of 100 bp
* 52694 56635: contig of 3942 bp in length
* 56636 56735: gap of 100 bp
* 56736 59688: contig of 2951 bp in length
* 59689 59786: gap of 100 bp
* 59787 63010: contig of 3224 bp in length
* 63011 63110: gap of 100 bp
* 63111 66323: contig of 3213 bp in length
* 66324 66423: gap of 100 bp
* 66424 69940: contig of 3517 bp in length
* 69941 70040: gap of 100 bp
* 70041 74249: contig of 4209 bp in length
* 74250 74349: gap of 100 bp
* 74350 77988: contig of 3639 bp in length
* 77989 82061: contig of 3973 bp in length
* 82062 82161: gap of 100 bp
* 82162 87385: contig of 5224 bp in length
* 87386 87485: gap of 100 bp
* 87486 92662: contig of 5177 bp in length
* 92663 92762: gap of 100 bp
* 92763 99121: contig of 6359 bp in length

```

```

* 99122 99221: gap of 100 bp
* 99222 104889: contig of 5668 bp in length
* 104890 104989: gap of 100 bp
* 104990 110282: contig of 5293 bp in length
* 110283 110382: gap of 100 bp
* 110383 118167: contig of 7785 bp in length
* 118168 118267: gap of 100 bp
* 118268 126038: contig of 7771 bp in length
* 126039 126138: gap of 100 bp
* 126139 132808: contig of 6670 bp in length
* 132809 132908: gap of 100 bp
* 132909 133337: contig of 429 bp in length.
FEATURES
    source
        1..133337
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_RP11="27H13"
            /clone_lib="RPC1-11 Human Male BAC"
            1..672
                /note="assembly_fragment"
            773..1578
                /note="assembly_fragment"
            1679..13396
                /note="assembly_fragment"
            13497..14553
                /note="assembly_fragment"
            14654..15808
                /note="assembly_fragment"
            15909..17297
                /note="assembly_fragment"
            17398..18345
                /note="assembly_fragment"
            18446..19764
                /note="assembly_fragment"
            19865..21198
                /note="assembly_fragment"
            21299..22590
                /note="assembly_fragment"
            22691..24188
                /note="assembly_fragment"
            24289..25165
                /note="assembly_fragment"
            25266..27286
                /note="assembly_fragment"
            27387..29216
                /note="assembly_fragment"
            29317..32525
                /note="assembly_fragment"
            32626..34252
                /note="assembly_fragment"
            34353..37416
                /note="assembly_fragment"
            37517..40514
                /note="assembly_fragment"
            40615..44387
                /note="assembly_fragment"
            44488..46779
                /note="assembly_fragment"
            46880..49231
                /note="assembly_fragment"
            49332..52593
                /note="assembly_fragment"
            52694..56635
                /note="assembly_fragment"
            56736..59686
                /note="assembly_fragment"
            59787..63010
                /note="assembly_fragment"
            63111..66323
                /note="assembly_fragment"
            66424..69940
                /note="assembly_fragment"

```

Query Match

72.0%; Score 18; DB 2; Length 133337;


```
repeat_region /rpt_family="(TA)n"
9392..9468
repeat_region /rpt_family="(TA)n"
9469..9513
repeat_region /rpt_family="LTR66"
complement(9524..9634)
/rpt_family="L2"
repeat_region 10482..10516
/rpt_family="AT-rich"
11445..11514
/rpt_family="GA-rich"
complement(11932..12116)
/rpt_family="L2"
repeat_region complement(13155..13416)
/rpt_family="MIR"
complement(15787..16085)
/rpt_family="Tigger3b"
16086..16398
/rpt_family="AluSx"
complement(16399..16946)
/rpt_family="Tigger3b"
16947..17255
/rpt_family="AluY"
complement(17256..17559)
/rpt_family="Tigger3b"
complement(17697..17769)
/rpt_family="MADE1"
18396..18512
/rpt_family="L2"
repeat_region 18609..18759
/rpt_family="GA-rich"
19259..19285
/rpt_family="(A)n"
19352..19385
/rpt_family="AluSx"
20538..20851
complement(21618..21792)
/rpt_family="MER5A"
complement(21940..22136)
/rpt_family="MIR"
22636..22747
/rpt_family="(TA)n"
complement(22848..23147)
/rpt_family="AluSx"
23531..23880
/rpt_family="MLR1J2"
complement(24450..24549)
/rpt_family="L2"
24884..25431
/rpt_family="L1PB1"
25436..27472
/rpt_family="L1PB1"
27473..27506
/rpt_family="(CAAA)n"
27507..28602
/rpt_family="L1PB1"
28603..28633
/rpt_family="(CAAA)n"
28634..31220
/rpt_family="L1PB1"
31221..31269
/rpt_family="(TG)n"
31264..31342
```

Query Match 72.0%; Score 18; DB 9; Length 146310;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTCCAGGAAACAGAAA 18
|||||
DB 51137 TGTCCAGGAAACAGAAA 51120

RESULT 13
AC111142/c
LOCUS
DEFINITION

AC111142 169013 bp DNA linear HTG 24-MAY-2002
Mus musculus clone RP23-370L24, WORKING DRAFT SEQUENCE, 7 ordered
pieces.

AC111142
VERSION AC111142.3 GI:21166275
HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
SOURCE Mus musculus
ORGANISM Mus musculus

REFERENCE
1 (bases 1 to 169013)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Mus musculus, clone RP23-370L24
Unpublished

2 (bases 1 to 169013)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choepe, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A.,
Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C.,
Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T.,
Mlonga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S.,
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (18-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 169013)

REFERENCE
AUTHORS

TITLE
JOURNAL

REFERENCE
AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L.,
Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepe, Y., Colangelo, M., Collins, S., Collamore, A.,
Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., FitzGerald, M., FitzHugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K.,
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K.,
Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N.,
Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J.,
Meneus, L., Mihova, T., Mlonga, V., Murphy, T., Naylor, J., Nguyen, C.,
Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C.,
Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S.,
Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S.,
Theodore, J., Topham, K., Travers, M., Travis, N., Trigglio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (24-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

On May 24, 2002 this sequence version replaced gi:21070739.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

DEFINITION Homo sapiens Chromosome 18 clone RP11-810C2 map 18q12, WORKING
DRAFT SEQUENCE, 13 unordered pieces.
ACCESSION AP001779
VERSION AP001779.3 GI:9188516
KEYWORDS HIG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens DNA, clone:RP11-810C2.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 191197)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 191,197 genomic DNA of 18q12
Published Only in DataBase (2000)
REFERENCE 2 (bases 1 to 191197)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (07-APR-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/
Tel:81-42-778-9923, Fax:81-42-778-9924)
On Jul 14, 2000 this sequence version replaced gi:8117451.
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDrafl18
Center clone name: RP11-810C2
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 186784 bases at least Q40
Consensus quality: 188417 bases at least Q30
Consensus quality: 189158 bases at least Q20
Insert size: 189997; sum-of-contigs
Quality coverage: 9.58x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
13 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
1 52211 contig of 52211 bp in length
52312 95959 contig of 43648 bp in length
96060 129191 contig of 33132 bp in length
129292 144917 contig of 15626 bp in length
145018 154847 contig of 9830 bp in length
154948 165265 contig of 10318 bp in length
165366 173461 contig of 8096 bp in length
173562 177237 contig of 3676 bp in length
177338 181320 contig of 3983 bp in length
181421 183769 contig of 2349 bp in length
183870 186900 contig of 3031 bp in length
187001 189395 contig of 2395 bp in length
189496 191197 contig of 1702 bp in length
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 52211: contig of 52211 bp in length
* 52212 52311: gap of 100 bp
* 52312 95959: contig of 43648 bp in length

95960 96059: gap of 100 bp
96060 129191: contig of 33132 bp in length
129192 129291: gap of 100 bp
129292 144917: contig of 15626 bp in length
144918 145017: gap of 100 bp
145018 154847: contig of 9830 bp in length
154848 154947: gap of 100 bp
154948 165265: contig of 10318 bp in length
165266 165365: gap of 100 bp
165366 173461: contig of 8096 bp in length
173462 173561: gap of 100 bp
173562 177237: contig of 3676 bp in length
177238 177337: gap of 100 bp
177338 181320: contig of 3983 bp in length
181321 181420: gap of 100 bp
181421 183769: contig of 2349 bp in length
183770 183869: gap of 100 bp
183870 186900: contig of 3031 bp in length
186901 187000: gap of 100 bp
187001 189395: contig of 2395 bp in length
189396 189495: gap of 100 bp
189496 191197: contig of 1702 bp in length.
FEATURES
Location/Qualifiers
1..191197
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18q12"
/clone="RP11-810C2"
1..52211
/note="assembly_fragment"
52312..95959
/note="assembly_fragment clone_end:77 vector_side:right"
96060..129191
/note="assembly_fragment"
129292..144917
/note="assembly_fragment"
145018..154847
/note="assembly_fragment"
154948..165265
/note="assembly_fragment"
165366..173461
/note="assembly_fragment"
173562..177237
/note="assembly_fragment clone_end:SP6 vector_side:left"
177338..181320
/note="assembly_fragment"
181421..183769
/note="assembly_fragment"
183870..186900
/note="assembly_fragment"
187001..189395
/note="assembly_fragment"
189496..191197
/note="assembly_fragment"
BASE COUNT 59669 a 35606 c 35044 g 59678 t 1200 others
ORIGIN
Query Match 72.0%; Score 18; DB 2; Length 191197;
Best Local Similarity 100.0%; Pred.No. 2.5;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TGTGCCAGGACACAGAAA 18
|||||
Db 111130 TGTGCCAGGACACAGAAA 111147
RESULT 18
CPA277161
LOCUS CPA277161 1433 bp DNA linear PLN 04-APR-2000
DEFINITION Carica papaya partial paccs1B gene for
1-aminocyclopropane-1-carboxylate synthase, exons 1-4.
ACCESSION AJ277161

AJ277161.1 GI:7413855
 1-aminocyclopropane-1-carboxylate synthase; paccs1b gene.
 Carica papaya
 ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Caricaceae; Carica.
 REFERENCE
 1 (bases 1 to 1433)
 Lam, P.F. and Abubakar, U.K.
 Nucleotide sequence of a partial DNA (PACCs1B) encoding ACC
 synthase from ripe papaya fruit
 JOURNAL
 Unpublished
 REFERENCE
 2 (bases 1 to 1433)
 Lam, P.F.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (03-APR-2000) Lam P.F., Biotechnology Programme,
 Malaysian Agricultural Research and Development Institute, GPO Box
 12301, 50774 Kuala Lumpur, MALAYSIA
 FEATURES
 source
 1..1433
 /organism="Carica papaya"
 /variety="Eksotika"
 /db_xref="taxon:3649"
 /clone="pPACCs1B"
 /tissue_type="ripe fruit"
 1..1433
 /gene="paccs1B"
 join(<1..24,136..264,374..535,649..>1433)
 /gene="paccs1B"
 /citation=[1]
 join(<1..24,136..264,374..535,649..>1433)
 /gene="paccs1B"
 /EC_number="4.4.1.14"
 /function="Converts SAM to ACC; plant hormone
 biosynthesis"
 /citation=[1]
 /codon_start=1
 /product="1-aminocyclopropane-1-carboxylate synthase"
 /protein_id="CAB86187.1"
 /db_xref="GI:7413856"
 /translation="OMGLAENQVSFLLEYLEQHSASTWGKAPRENALFDYH
 GLKSFQAMSEFIQIRGKADPERIVLTAGATAANELLFIILNGDALLVPTPY
 YPGFDRDLRWRTGKIVIPHCDSNNFQITPEALSAINEASMMVRGILITPSN
 PLGITIPRVLQDLDFVTLKTHLISDETSQSVPTSDFTSVVAELHRRNNVAECN
 ERVHIVSLKDLGLPGFRVGTIYNKVTITARMSSFTLIFSQTQHLHLSMLSNQ
 KFTENVINTNRDLRRRYENIVEGLKRSAGIECLAGNAGLCFWMNLSSFLDEPTREAEI
 NLNDSILHRVKNLSPGSSCHSCSEPGWFRVQFA"
 <1..24
 /gene="paccs1B"
 /number=1
 25..135
 /gene="paccs1B"
 /number=1
 136..264
 /gene="paccs1B"
 /number=2
 265..373
 /gene="paccs1B"
 /number=2
 374..535
 /gene="paccs1B"
 /number=3
 536..648
 /gene="paccs1B"
 /number=3
 649..>1433
 /gene="paccs1B"
 /number=4
 BASE COUNT 424 a 296 c 282 g 431 t
 ORIGIN
 Query Match 68.0%; Score 17; DB 8; Length 1433;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 GAACACAGAAATTACAG 25
 |||||
 DB 1155 GAACACAGAAATTACAG 1171
 |||||
 RESULT 19
 AC027349 40740 bp DNA linear PRI 06-AUG-2002
 LOCUS Homo sapiens chromosome 16 clone CTD-3083H4, complete sequence.
 DEFINITION
 AC027349
 ACCESSION
 VERSION AC027349.3 GI:22122874
 KEYWORDS
 SOURCE HTG.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 40740)
 DOE Joint Genome Institute.
 TITLE Sequencing of Human Chromosome 16
 JOURNAL Unpublished
 REFERENCE
 2 (bases 1 to 40740)
 DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (30-MAR-2000) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 3 (bases 1 to 40740)
 DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (06-AUG-2002) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 On Aug 6, 2002 this sequence version replaced gi:9090640.
 COMMENT
 Sequence Quality Assessment:
 This entry has been annotated with sequence quality
 estimates computed by the Phrap assembly program.
 All manually edited bases have been reduced to quality zero.
 Quality levels above 40 are expected to have less than
 1 error in 10,000 bp.
 Base-by-base quality values are not generally visible from the
 Genbank flat file format but are available as part
 of this entry's ASN.1 file.

 Sequence Quality Assessment:
 This entry has been annotated with sequence quality
 estimates computed by the Phrap assembly program.
 All manually edited bases have been reduced to quality zero.
 Quality levels above 40 are expected to have less than
 1 error in 10,000 bp.
 Base-by-base quality values are not generally visible from the
 Genbank flat file format but are available as part
 of this entry's ASN.1 file.

 FEATURES
 source
 Location/Qualifiers
 1..40740
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="16"
 /clone="CTD-3083H4"
 BASE COUNT 11000 a 9214 c 10026 g 10500 t
 ORIGIN
 Query Match 68.0%; Score 17; DB 9; Length 40740;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 TGCACAGAACACAGAAAT 19
 |||||
 DB 40194 TGCACAGAACACAGAAAT 40210
 |||||
 RESULT 20
 AC119489_3

WPCOMMENT

Sequence split into 4 fragments LOCUS AC119489 Accession AC119489
Fragment Name Begin End
AC119489_0 1 110000
AC119489_1 100001 210000
AC119489_2 200001 310000
AC119489_3 300001 35252

Continuation (4 of 4) of AC119489 from base 300001 (AC119489 Rattus norvegicus clone CH2)

Query Match 68.0%; Score 17; DB 2; Length 55252;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GCCAGGACCCAGAAATT 20
|||||||

Db 54377 GCCAGGACCCAGAAATT 54393

RESULT 21

AC128098/c

LOCUS AC128098 104934 bp DNA linear HTG 19-JUL-2002
DEFINITION Rattus norvegicus clone CH230-322L19, *** SEQUENCING IN PROGRESS
***, 54 unordered pieces.

ACCESSION

AC128098.1 GI:21908696

VERSION HTG: HTGS PHASEL.

KEYWORDS Rattus norvegicus.

SOURCE

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

1 (bases 1 to 104934)

AUTHORS

Murphy, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,
Alsbrooks, S.L., Amarantunga, H.C., Are, J.R., Ayele, M., Banks, T.,
Barack, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, K.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorelli, H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hoques, M., Holloway, C., Hollins, B.,
Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jollivet, S., Joudah, S.,
Karlssohn, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,
Lorado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenko, S., Ogih, M., Okwuonu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I.,
Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tayor, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, I., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.

TITLE

Direct Submission

JOURNAL

REFERENCE 2 (bases 1 to 104934)

AUTHORS

Worley, K.C.

TITLE

Direct Submission

JOURNAL

Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GZOM
Center clone name: CH230-322L19
----- Summary Statistics
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 5882 bases at least Q40
Consensus quality: 64470 bases at least Q30
Consensus quality: 68220 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 54 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 1190: contig of 1190 bp in length
* 1191 1290: gap of unknown length
* 1291 2303: contig of 1013 bp in length
* 2304 2403: gap of unknown length
* 2404 3435: contig of 1032 bp in length
* 3436 3535: gap of unknown length
* 3536 4670: contig of 1135 bp in length
* 4671 4770: gap of unknown length
* 4771 6482: contig of 1712 bp in length
* 6483 6582: gap of unknown length
* 6583 7869: contig of 1287 bp in length
* 7870 7969: gap of unknown length
* 7970 9257: contig of 1288 bp in length
* 9258 9357: gap of unknown length
* 9358 10734: contig of 1377 bp in length
* 10735 10834: gap of unknown length
* 10835 11899: contig of 1065 bp in length
* 11900 11999: gap of unknown length
* 12000 13222: contig of 1223 bp in length
* 13223 13322: gap of unknown length
* 13323 14810: contig of 1488 bp in length
* 14811 16075: contig of 1165 bp in length
* 16076 16175: gap of unknown length
* 16176 17275: contig of 1100 bp in length
* 17276 17375: gap of unknown length
* 17376 18935: contig of 1460 bp in length
* 18936 20535: contig of 1600 bp in length
* 20536 22175: contig of 1540 bp in length
* 22176 22275: gap of unknown length
* 22276 24279: contig of 2004 bp in length
* 24280 25506: contig of 1127 bp in length
* 25507 25606: gap of unknown length
* 25607 27014: contig of 1408 bp in length
* 27015 27114: gap of unknown length
* 27115 28993: contig of 1879 bp in length
* 28994 29093: gap of unknown length

* 29094 30571: contig of 1478 bp in length
* 30572 30671: gap of unknown length
* 30672 32327: contig of 1656 bp in length
* 32328 32427: gap of unknown length
* 32428 33824: contig of 1397 bp in length
* 33825 33924: gap of unknown length
* 33925 35527: contig of 1603 bp in length
* 35528 37444: contig of 2117 bp in length
* 37445 37844: gap of unknown length
* 37845 39102: contig of 1258 bp in length
* 39103 39202: gap of unknown length
* 39203 40775: contig of 1573 bp in length
* 40776 42622: contig of 1747 bp in length
* 42623 42722: gap of unknown length
* 42723 43932: contig of 1210 bp in length
* 43933 44032: gap of unknown length
* 44033 45667: contig of 1635 bp in length
* 45668 45767: gap of unknown length
* 47355 47354: contig of 1587 bp in length
* 47355 47454: gap of unknown length
* 47455 48993: contig of 1539 bp in length
* 48994 49093: gap of unknown length
* 49094 50421: contig of 1328 bp in length
* 50422 50521: gap of unknown length
* 50522 52612: contig of 2091 bp in length
* 52613 52712: gap of unknown length
* 52713 54772: contig of 2060 bp in length
* 54773 54872: gap of unknown length
* 54873 56320: contig of 1448 bp in length
* 56321 56420: gap of unknown length
* 56421 58547: contig of 2127 bp in length
* 58548 58647: gap of unknown length
* 58648 60811: contig of 2164 bp in length
* 60812 60911: gap of unknown length
* 60912 62214: contig of 1303 bp in length
* 62215 62314: gap of unknown length
* 62315 64232: contig of 1918 bp in length
* 64233 64332: gap of unknown length
* 64333 67357: contig of 3025 bp in length
* 67358 67457: gap of unknown length
* 67458 70118: contig of 2661 bp in length
* 70119 70218: gap of unknown length
* 70219 71878: contig of 1660 bp in length
* 71879 71978: gap of unknown length
* 71979 73325: contig of 1347 bp in length
* 73326 73425: gap of unknown length
* 73426 75179: contig of 1754 bp in length
* 75180 75279: gap of unknown length
* 75280 77923: contig of 2644 bp in length
* 77924 78023: gap of unknown length
* 78024 79934: contig of 1911 bp in length
* 79935 80034: gap of unknown length
* 80035 83274: contig of 3240 bp in length
* 83275 83374: gap of unknown length
* 83375 86522: contig of 3148 bp in length
* 86523 86623: gap of unknown length
* 86623 89039: contig of 2417 bp in length
* 89040 89139: gap of unknown length
* 89140 94471: contig of 5332 bp in length
* 94472 94571: gap of unknown length
* 94572 97846: contig of 3275 bp in length
* 97847 97946: gap of unknown length
* 97947 102259: contig of 4313 bp in length
* 102260 102359: gap of unknown length
* 102360 104934: contig of 2575 bp in length.

FEATURES

source

Location/Qualifiers
1. .104934
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-322L19"

BASE COUNT

30320 a 19586 c 20195 g 29489 t 5344 others

ORIGIN

Query Match 68.0%; Score 17; DB 2; Length 104934;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AGGAACACGAAATTTAC 23
|||||
Db 42401 AGGAACACGAAATTTAC 42385

RESULT 22

AC118386

LOCUS

DEFINITION

AC118386

VERSION

AC118386.2 GI:21746308

KEYWORDS

HTG; HTGS_PHASE1.

SOURCE

Norway rat.

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE

1 (bases 1 to 107361)

AUTHORS

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,

Alsbrooks, S.L., Amarantunga, H.C., Are, J.R., Ayale, M., Banks, T.,

Barbieri, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D.,

Bouck, J., Bowler, S., Brieva, M., Brown, M., Brown, N.P.,

Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,

Carroll, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, C.,

Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,

Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,

Devila, M.L., Davis, C., Davy Carroll, L., Dederich, D.A.,

Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,

Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,

Einhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,

Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,

Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,

Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,

Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,

Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,

Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,

Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,

Karlsone, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,

Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,

Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H.,

Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,

Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,

Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,

Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,

Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,

Nguyen, N., Nickerson, E., Nwokenwo, S., Ogih, M., Okwunigbo, G.,

Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,

Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,

Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G.,

Schere, S., Scott, G., Shen, H., Shoshitari, N., Sisson, I.,

Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,

Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,

Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,

Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,

Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,

Williams, G., Williamson, A., Wleczek, R., Wleczek, S., Wolek, K.,

Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,

Weinstock, G. and Gibbs, R.

TITLE

JOURNAL

REFERENCE

2 (bases 1 to 107361)

Unpublished

Worley, K.C.

Direct Submission

Submitted (17-APR-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
TITLE
JOURNAL

3 (bases 1 to 107361)
Morley K.C.
Direct Submission
Submitted (18-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 14, 2002 this sequence version replaced gi:20162749.

COMMENT

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GUEX
Center clone name: CH230-397N16
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 65518 bases at least Q40
Consensus quality: 68520 bases at least Q30
Consensus quality: 70855 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: this is a 'working draft' sequence. It currently
* consists of 47 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
1201: contig of 1201 bp in length
1301: gap of unknown length
1302: contig of 1125 bp in length
2427: gap of unknown length
2527: gap of unknown length
3672: contig of 1146 bp in length
3673: gap of unknown length
3773: contig of 1124 bp in length
4897: gap of unknown length
6187: contig of 1191 bp in length
6188: gap of unknown length
6288: contig of 1131 bp in length
7419: gap of unknown length
7519: contig of 1446 bp in length
8965: gap of unknown length
9065: contig of 1079 bp in length
10143: gap of unknown length
10243: gap of unknown length
10244: contig of 1368 bp in length
11611: gap of unknown length
11612: contig of 1197 bp in length
11711: gap of unknown length
12908: contig of 1197 bp in length
12909: gap of unknown length
13008: gap of unknown length
14201: contig of 1193 bp in length
14301: gap of unknown length
14302: contig of 1241 bp in length
15542: gap of unknown length
15642: gap of unknown length
17546: contig of 1904 bp in length
17547: gap of unknown length
17647: contig of 1703 bp in length
19349: gap of unknown length
19350: gap of unknown length
20867: contig of 1418 bp in length
20868: gap of unknown length
20957: gap of unknown length
22520: contig of 1553 bp in length
22620: gap of unknown length
24029: contig of 1409 bp in length
24030: gap of unknown length
24129: gap of unknown length
25659: contig of 1530 bp in length
25759: gap of unknown length
27750: contig of 1991 bp in length
27751: gap of unknown length
28914: contig of 1064 bp in length
29014: gap of unknown length
28915

29015
31201: contig of 2187 bp in length
31202: gap of unknown length
31301: contig of 1258 bp in length
32559: gap of unknown length
32560: contig of 1022 bp in length
33681: gap of unknown length
33682: contig of 1683 bp in length
33781: gap of unknown length
35464: gap of unknown length
35465: contig of 1380 bp in length
36944: gap of unknown length
36945: contig of 2602 bp in length
37045: gap of unknown length
39647: contig of 2034 bp in length
39747: gap of unknown length
41781: contig of 2511 bp in length
41880: gap of unknown length
41881: contig of 2109 bp in length
44391: gap of unknown length
44392: contig of 2109 bp in length
44491: gap of unknown length
44492: contig of 2109 bp in length
46601: gap of unknown length
46700: contig of 1430 bp in length
48130: gap of unknown length
48230: contig of 3035 bp in length
48231: gap of unknown length
51265: gap of unknown length
51266: contig of 3571 bp in length
51363: gap of unknown length
51366: contig of 2596 bp in length
54937: gap of unknown length
55036: contig of 2170 bp in length
55037: gap of unknown length
57632: gap of unknown length
57633: contig of 2227 bp in length
59002: gap of unknown length
59003: contig of 1810 bp in length
60002: gap of unknown length
60003: contig of 2956 bp in length
62230: gap of unknown length
62239: gap of unknown length
64139: contig of 2956 bp in length
64140: gap of unknown length
64239: gap of unknown length
67195: gap of unknown length
67196: contig of 4133 bp in length
67295: gap of unknown length
67296: contig of 2506 bp in length
71428: gap of unknown length
71429: gap of unknown length
71529: contig of 3001 bp in length
74035: gap of unknown length
74134: gap of unknown length
77135: contig of 4356 bp in length
77136: gap of unknown length
77236: contig of 3338 bp in length
81591: gap of unknown length
81592: contig of 3431 bp in length
81692: gap of unknown length
85122: gap of unknown length
85123: contig of 3338 bp in length
85223: gap of unknown length
88561: gap of unknown length
88562: contig of 4236 bp in length
92896: gap of unknown length
92897: contig of 4298 bp in length
92997: gap of unknown length
97295: contig of 4964 bp in length
97394: gap of unknown length
97395: contig of 4964 bp in length
102359: gap of unknown length
102459: contig of 4903 bp in length.

FEATURES
source

1. 107361
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-397N16"

BASE COUNT 26683 a 21519 c 21208 g 26730 t 11221 others
ORIGIN

Query Match 68.0%; Score 17; DB 2: Length 107361;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GCCAGGACCCAGAAATT 20

|||||

Db 9135 GCCAGGACCCAGAAATT 9151

RESULT 23

CEYL11B2_2

WPCOMMENT

Sequence split into 7 fragments LOCUS CEY11B2 Accession Z98857

Fragment Name Begin End
CEY11B2_0 1 110000
CEY11B2_1 100001 210000
CEY11B2_2 200001 310000
CEY11B2_3 300001 410000
CEY11B2_4 400001 510000
CEY11B2_5 500001 610000
CEY11B2_6 600001 614429

Continuation (3 of 7) of CEY11B2 from base 200001 (Z98857 Caenorhabditis elegans chromosome

Query Match 68.0%; Score 17; DB 2; Length 110000;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CAGGAACCAAGAAATTTA 22

Db 74352 CAGGAACCAAGAAATTTA 74368

RESULT 24
AL365204/c

LOCUS Human DNA sequence from clone Rpl1-321L2 on chromosome 9, complete
DEFINITION 129191 bp DNA linear PRI 24-APR-2001

ACCESSION AL365204 AC073983
VERSION AL365204.11 GI:13810068
KEYWORDS HTG.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 129191)
Direct Submission
Submitted (24-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonesrequest@sanger.ac.uk

On or before May 15, 2001 this sequence version replaced
gi:8979993, gi:13396592.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em.; EMBL; SW.;
SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr9>

COMMENT

Rpl1-321L2 is from the library RPl1-11.2 constructed by the group
of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
VECTOR: pBACe3.6
IMPORTANT: This sequence is not the entire insert of clone
Rpl1-321L2. It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true right end of clone Rpl1-321L2 is at 129191 in this
sequence. The true right end of clone Rpl1-31K16 is at 100 in this
sequence.

FEATURES

Location/Qualifiers

source
1..129191
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="Rpl1-321L2"
/clone_lib="RPCI-11.2"
1262..1428
/note="MIR repeat: matches 89..262 of consensus"
3496..3539
/note="11 copies 4 mer tgtg 84% conserved"
4371..4499
/note="MIR repeat: matches 11..146 of consensus"
6913..7101
/note="MIR repeat: matches 56..250 of consensus"
7125..7300
/note="MER5B repeat: matches 6..178 of consensus"
10943..11213
/note="L2 repeat: matches 2460..2705 of consensus"
17072..17119
/note="16 copies 3 mer acc 79% conserved"
17572..17659
/note="44 copies 2 mer tc 88% conserved"
17574..17657
/note="21 copies 4 mer tctc 90% conserved"
17688..17717
/note="15 copies 2 mer ca 86% conserved"
17765..18376
/note="CpG island"
/evidence="not_experimental"
19301..19596
/note="AluY repeat: matches 1..295 of consensus"
21560..21624
/note="MLT1J repeat: matches 109..173 of consensus"
21629..21925
/note="AluX repeat: matches 1..296 of consensus"
22007..22263
/note="MLT1J repeat: matches 228..482 of consensus"
22749..22904
/note="FRAM repeat: matches 7..157 of consensus"
23487..23712
/note="MIR repeat: matches 29..261 of consensus"
24265..24320
/note="MIR repeat: matches 88..144 of consensus"
24725..24971
/note="AluY repeat: matches 1..297 of consensus"
25030..25216
/note="MER5A repeat: matches 1..189 of consensus"
25239..25622
/note="MLT1H repeat: matches 52..503 of consensus"
26274..26572
/note="AluY repeat: matches 1..299 of consensus"
27476..33907
/note="LIP7 repeat: matches 7..6141 of consensus"
34090..34561
/note="MLT1D repeat: matches 1..493 of consensus"
36528..40748
/note="LIP2 repeat: matches 1814..6155 of consensus"
40886..41065
/note="LIMB8 repeat: matches 5947..6140 of consensus"
41064..41123
/note="LIM4 repeat: matches 5667..5727 of consensus"
41344..41405
/note="MIR repeat: matches 3..59 of consensus"
41406..41759
/note="THELC repeat: matches 1..370 of consensus"
41760..41892
/note="MIR repeat: matches 59..212 of consensus"
41948..42777
/note="LIMB2 repeat: matches 5318..6165 of consensus"
42795..44704
/note="LIP2 repeat: matches 4140..6155 of consensus"
44709..44898
/note="L1 repeat: matches 5144..5330 of consensus"

repeat_region 46070..48455
/note="L1M45A repeat: matches 3961..15285 of consensus"
repeat_region 48474..48835
/note="THE1B repeat: matches 1..364 of consensus"
repeat_region 48836..50419
/note="THE1B-INTERNAL repeat: matches 1..1580 of consensus"
repeat_region 50420..50492
/note="THE1C repeat: matches 300..371 of consensus"
repeat_region 50514..51080
/note="L1M1 repeat: matches 1134..1658 of consensus"
repeat_region 51082..51578
/note="MLT2CB repeat: matches 12..501 of consensus"
repeat_region 51712..52013
/note="AluJo repeat: matches 1..296 of consensus"
repeat_region 52127..52211
/note="AluJ/FLAM repeat: matches 1..85 of consensus"
repeat_region 52315..52473
/note="FRAM repeat: matches -6..146 of consensus"
repeat_region 52492..52773
/note="AluX repeat: matches 1..282 of consensus"
repeat_region 52786..52819
/note="17 copies 2 mer at 85% conserved"
repeat_region 52788..52819
/note="8 copies 4 mer at 87% conserved"
repeat_region 53063..53888
/note="L1M4C repeat: matches 1242..1587 of consensus"
repeat_region 53960..54096
/note="L1M4C repeat: matches 1009..1143 of consensus"
repeat_region 54097..54513
/note="MSTA repeat: matches 1..426 of consensus"
repeat_region 54514..54596
/note="L1M4C repeat: matches 926..1009 of consensus"
repeat_region 55756..56074
/note="L1M4 repeat: matches 4154..4470 of consensus"
repeat_region 56084..56452
/note="L2 repeat: matches 2373..2747 of consensus"
repeat_region 56972..57204
/note="L1MD2 repeat: matches 5778..6001 of consensus"
repeat_region 57687..57976
/note="AluX repeat: matches 1..290 of consensus"
repeat_region 58330..59097
/note="L1M4A repeat: matches 6023..6295 of consensus"
repeat_region 59105..59416
/note="AluSq repeat: matches 1..313 of consensus"
repeat_region 59811..59910
/note="AluY repeat: matches 211..310 of consensus"
repeat_region 60755..62450
/note="L1FA5 repeat: matches 4445..6143 of consensus"
repeat_region 62558..62765
/note="MIR repeat: matches 52..260 of consensus"
repeat_region 64608..64706
/note="L1MC5 repeat: matches 7835..7939 of consensus"
repeat_region 64930..65232
/note="AluY repeat: matches 3..310 of consensus"
repeat_region 65234..65595
/note="MER57B repeat: matches 1..390 of consensus"
repeat_region 65596..66194
/note="L1TR29 repeat: matches 4..616 of consensus"
repeat_region 66277..66588
/note="L1ME3 repeat: matches 5655..5988 of consensus"
repeat_region 67535..67638
/note="MIR repeat: matches 47..150 of consensus"
repeat_region 67727..67925
/note="AluJo repeat: matches 83..309 of consensus"
repeat_region 69066..69174
/note="FLAM_A repeat: matches 1..131 of consensus"
repeat_region 69458..70040
/note="MLT2D repeat: matches 2..553 of consensus"
repeat_region 71046..71085
/note="MER5A repeat: matches 128..167 of consensus"
repeat_region 71047..71111
/note="MER5A repeat: matches 35..107 of consensus"

repeat_region 71368..71686
/note="MLT2E repeat: matches 3..330 of consensus"
repeat_region 71687..71848
Query Match 68.0%; Score 17; DB 9; Length 129191;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 TGCCAGGAGACACAGAAAT 19
|||||
Db 120038 TCCAGGAGACACAGAAAT 120022
RESULT 25
AL356862/c 145481 bp DNA linear PRI 19-JAN-2001
LOCUS Human DNA sequence from clone RP11-1M19 on chromosome q34.11-34.3,
DEFINITION complete sequence.
ACCESSION AL356862
VERSION AL356862.10 GI:12044598
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 145481)
AUTHORS Garner P.
TITLE Direct Submission
JOURNAL Submitted (19-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
On Jan 8, 2001 this sequence version replaced gi:11611012.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated repeat sequence elements. Where the sequence is
ambiguous, there is an annotation using the 'unsure' feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em.: EMBL; Sw.: SWISSPROT; Tr.: TREMBL; Wp.: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C.elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9
RP11-1M19 is from the library RPC1-11.1 constructed by the group of
Pietter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
This sequence is the entire insert of clone RP11-1M19 The true left
end of clone RP11-342E24 is at 59483 in this sequence. The true
right end of clone RP11-405c6 is at 27333 in this sequence.

FEATURES
source
1..145481
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/map="q34.11-34.3"
/clone="RP11-1M19"
/clone_lib="RPC1-11.1"

BASE COUNT 36669 a 33113 c 34164 g 41535 t
ORIGIN

Query Match 68.0%; Score 17; DB 9; Length 145481;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      4  GCCAGGACCAAGAAAT 20
        |||||
Db 77828 GCCAGGACCAAGAAAT 77812

RESULT 26
AC018586/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

    160538 bp      DNA      linear      HTG 12-MAR-2000
Homo sapiens clone RP11-9L23, WORKING DRAFT SEQUENCE, 7 unordered
pieces
AC018586
AC018586.3 GI:7230039
HTG; HTGS_PHASE1; HTGS_DRAFT.
Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 160538)
Birten,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguslavsky,L., Bouckghalter,B., Brown,A., Burkett,G., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeRellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lechoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
Meldrum,J., Meneus,L., Morrow,J., Naylor,J., Norman,C.H.,
O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A.,
Santos,K., Severy,P., Spencer,B., Stange-Thomann,N., Stofanovic,N.,
Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
Vassiliev,H., Viel,R., Vo.A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A.
and Zody,M.
Direct Submission
Submitted (14-DEC-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 12, 2000 this sequence version replaced gi:6648257.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3018
Center clone name: 9_L23
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 150680 bases at least Q40
Consensus quality: 156181 bases at least Q30
Consensus quality: 158397 bases at least Q20
Insert size: 165000; agarose-fp
Quality coverage: 4.1 in Q20 bases; sum-of-contigs
Quality coverage: 4.2 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

```

as soon as it is available and the accession number will
* be preserved.

1 4906: contig of 4906 bp in length
* 4907 5006: gap of 100 bp
* 12883: contig of 7877 bp in length
* 12884 12983: gap of 100 bp
* 12984 29048: contig of 16065 bp in length
* 29049 29148: gap of 100 bp
* 29149 59139: contig of 29991 bp in length
* 59140 59239: gap of 100 bp
* 59240 89099: contig of 29860 bp in length
* 89100 89199: gap of 100 bp
* 89200 121942: contig of 32743 bp in length
* 121943 122042: gap of 100 bp
* 122043 160538: contig of 38496 bp in length.

FEATURES
 source
 1..160538
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="RP11-9L23"
 /clone_lib="RPC1-11 Human Male BAC"
 misc_feature
 1..4906
 /note="assembly_fragment"
 misc_feature
 5007..12883
 /note="assembly_fragment"
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:right
 12984..29048
 /note="assembly_fragment"
 misc_feature
 29149..59139
 /note="assembly_fragment"
 /note="assembly_fragment"
 clone_end:T7
 vector_side:right
 59240..89099
 /note="assembly_fragment"
 misc_feature
 89200..121942
 /note="assembly_fragment"
 misc_feature
 122043..160538
 /note="assembly_fragment"

BASE COUNT 47831 a 29592 c 29049 g 53460 t 606 others
 ORIGIN

Query Match 68.08; Score 17; DB 2; Length 160538;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TGCCAGGACCAAGAAAT 19
 |||||
 Db 92137 TGCCAGGACCAAGAAAT 92121

RESULT 27
 AL355478/c
 LOCUS
 DEFINITION
 AL355478 Human DNA sequence from clone RP11-45J023 on chromosome 13,
 complete sequence.
 ACCESSION
 AL355478
 VERSION
 AL355478.16 GI:11876038
 KEYWORDS
 HTG.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 174736)
 Oliver,K.
 Direct Submission
 Submitted (15-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On Dec 17, 2000 this sequence version replaced gi:11544993.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated repeat sequence elements. Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em.; EMBL; Sw.; SWISSPROT; Tr.; TRMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at

<http://www.sanger.ac.uk/projects/C-elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human

chromosome 13, constructed by the Sanger Centre Chromosome 13 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr13>

RP11-459J23 is from the library RP11-11.2 constructed by the group of Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6

This sequence is the entire insert of clone RP11-459J23 The true left end of clone RP11-237E3 is at 120113 in this sequence. The true right end of clone RP11-501G6 is at 18764 in this sequence.

FEATURES

```

source
1..174736
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="13"
    /clone="RP11-459J23"
    /clone_lib="RP11-11.2"
    796..827
        /note="16 copies 2 mer to 87% conserved"
    853..880
        /note="14 copies 2 mer ac 89% conserved"
    4826..5290
        /note="L1MA7 repeat: matches 5824..5283 of consensus"
    5914..6070
        /note="L1MA7 repeat: matches 1994..2151 of consensus"
    9261..9399
        /note="L2 repeat: matches 1283..1426 of consensus"
    10111..10378
        /note="MIR repeat: matches 6..262 of consensus"
    12261..12405
        /note="MIR repeat: matches 59..214 of consensus"
    13124..13248
        /note="L2 repeat: matches 2371..2496 of consensus"
    13326..13467
        /note="L2 repeat: matches 2601..2750 of consensus"
    13949..13984
        /note="L2 copies 3 mer aa 86% conserved"
    15107..15170
        /note="32 copies 2 mer tg 93% conserved"
    15634..15806
        /note="FAM repeat: matches 1..167 of consensus"
    16438..16509
        /note="36 copies 2 mer aa 66% conserved"
    17539..17616
        /note="MIR repeat: matches 33..115 of consensus"
    18506..18533
        /note="14 copies 2 mer gt 92% conserved"
    19491..19546
        /note="28 copies 2 mer tt 69% conserved"
    20234..20520
        /note="AluSg repeat: matches 1..287 of consensus"
    20945..21234
        /note="AluY repeat: matches 2..307 of consensus"
    21679..21887
        /note="L1PA8 repeat: matches 5955..6163 of consensus"
    21888..22683
        /note="L1P repeat: matches 3673..4469 of consensus"
    23134..23182
        /note="HSMAR1 repeat: matches 1237..1286 of consensus"

```

```

repeat_region
28857..28916
    /note="2 copies 30 mer 93% conserved"
29207..29650
    /note="MER65A repeat: matches 1..445 of consensus"
29651..29680
    /note="15 copies 2 mer tt 86% conserved"
29692..31011
    /note="MER65-internal repeat: matches 3513..4851 of consensus"
31100..31412
    /note="MER31-internal repeat: matches 5305..5593 of consensus"
31413..31707
    /note="AluY repeat: matches 1..295 of consensus"
31708..32405
    /note="MER31-internal repeat: matches 3077..5305 of consensus"
32398..32706
    /note="MER4-internal repeat: matches 3734..4021 of consensus"
32534..32712
    /note="MER83-internal repeat: matches 3136..3309 of consensus"
32761..32880
    /note="MER83-internal repeat: matches 3017..3136 of consensus"
33026..33262
    /note="MER4-internal repeat: matches 3170..3395 of consensus"
33396..33529
    /note="FLAM_C repeat: matches 10..143 of consensus"
33779..33985
    /note="AluY repeat: matches 106..310 of consensus"
34176..35049
    /note="MER65-internal repeat: matches 1647..2494 of consensus"
35078..35358
    /note="L1MA7 repeat: matches 5998..6279 of consensus"
35388..35980
    /note="L1 repeat: matches 2445..3040 of consensus"
35995..36050
    /note="14 copies 4 mer tatc 85% conserved"
36111..36254
    /note="36 copies 4 mer atat 70% conserved"
36129..36248
    /note="4 copies 30 mer 73% conserved"
36259..36385
    /note="MSRD repeat: matches 260..392 of consensus"
36533..36679
    /note="MSTC repeat: matches 9..155 of consensus"
36705..38179
    /note="L1M3e repeat: matches 1222..2794 of consensus"
38180..38564
    /note="MSRD repeat: matches 1..394 of consensus"
38565..39381
    /note="L1M3e repeat: matches 474..1222 of consensus"
39987..40070
    /note="L2 repeat: matches 2114..2197 of consensus"
40217..41560
    /note="L2 repeat: matches 963..2332 of consensus"
41566..42604
    /note="L1PA7 repeat: matches 5099..6126 of consensus"
42636..42925
    /note="L2 repeat: matches 685..977 of consensus"
43503..43884
    /note="MIR repeat: matches 164..547 of consensus"
47044..47091
    /note="24 copies 2 mer ac 87% conserved"
48849..49072
    /note="112 copies 2 mer tt 55% conserved"
49159..49299
    /note="MIR repeat: matches 114..262 of consensus"
49388..49539

```



```

repeat_region /note="L2 repeat: matches 2264. .2405 of consensus"
51927. .52472
repeat_region /note="L1 repeat: matches 3473. .4016 of consensus"
52438. .53721
repeat_region /note="LIPAS repeat: matches 4862. .6143 of consensus"
53734. .54794
repeat_region /note="LIM1 repeat: matches 4453. .5403 of consensus"
54795. .55168
repeat_region /note="THE1C repeat: matches 1. .371 of consensus"
55169. .55420
repeat_region /note="LIPB3 repeat: matches 5537. .5791 of consensus"
55446. .55795
repeat_region /note="LIPB3 repeat: matches 5776. .6127 of consensus"
55810. .55847
repeat_region /note="19 copies 2 mer aa 81% conserved"
56194. .56486
repeat_region /note="ALUSx repeat: matches 1. .298 of consensus"
56520. .56875
repeat_region /note="L2 repeat: matches 2179. .2320 of consensus"
56912. .57251
repeat_region /note="L2 repeat: matches 1691. .2016 of consensus"
59460. .59607
repeat_region /note="LIME2 repeat: matches 5836. .5985 of consensus"
60584. .60886
repeat_region /note="LIM4 repeat: matches 3020. .3317 of consensus"
60931. .62103
repeat_region /note="LIPAL6 repeat: matches 1708. .2863 of consensus"
62104. .62395
repeat_region /note="ALUy repeat: matches 1. .290 of consensus"
62396. .62687
repeat_region /note="LIPAL6 repeat: matches 2863. .3157 of consensus"
62750. .65798
repeat_region /note="LIPAL6 repeat: matches 3178. .6155 of consensus"
67132. .67331

Query Match 68.0%; Score 17; DB 9; Length 174736;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CAGGAACCAAGAAATTTA 22
|||||
Db 116422 CAGGAACCAAGAAATTTA 116406

RESULT 28
AC097065
LOCUS AC097065 176861 bp DNA linear PRI 21-FEB-2002
DEFINITION Homo sapiens chromosome 1 clone RP11-532L16, complete sequence.
ACCESSION AC097065 AL358153
VERSION AC097065.2 GI:18855081
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 176861)
AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Raymond,C. and Haugen,E.D.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 176861)
AUTHORS Kaul,R.K., Olson,M.V., Raymond,C., Clendenning,J. and Haugen,E.D.
TITLE Direct Submission
JOURNAL Submitted (06-OCT-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
REFERENCE 3 (bases 1 to 176861)
AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Raymond,C. and Haugen,E.D.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
COMMENT On Feb 21, 2002 this sequence version replaced gi:15982508.
----- Genome Center

```

Center: University of Washington Genome Center
Center Code: UWGC
Web site: <http://www.genome.washington.edu>
Contact: uwgchgs@u.washington.edu
Drafting Center: SC

Project Information

Center project name: chr-1
Center clone name: RP11-532L16 (sc0733)
----- Summary Statistics -----
Sequencing vector: plasmid; 48% of reads
Chemistry: Dye-terminator ET; 58% of reads
Chemistry: Dye-terminator Big Dye; 42% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 176627 bases at least Q40
Consensus quality: 176846 bases at least Q30
Consensus quality: 176861 bases at least Q20
Insert size: 176495; sum-of-contigs
Quality coverage: 7.3x in Q20 bases; sum-of-contigs

Overlapping Sequences:

5': Mapping in progress
3': RP11-469A15 (UWGC:sc0718), AC104461

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

EcoRI			BglII			HindIII		
SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap
8696	8887	9143	8853	8208	7969			
6	<800	2067	2035	6382	6259			
261	<800	5161	5172	512	<800			
273	<800	10110	9769	449	<800			
10535	10457	730	<800	8459	8418			
532	<800	1480	1435	26	<800			
8513	8887	5215	5369	2487	2524			

3891	3859	455	<800	713	<800
2555	2542	3352	3310	2554	2524
12086	11736	5973	5896	18764	19124
11598	11291	2818	2829	17982	17787
5497	5533	5914	5896	2326	2375
3007	3017	1863	1877	1578	1565
7124	7208	1567	1526	9232	9104
642	<800	10015	9769	3804	3794
6549	6624	5405	5369	1080	1096
1561	1528	351	<800	166	<800
2485	2542	1018	1080	408	<800
1825	1788	22	<800	198	<800
3011	3017	1097	1080	10206	10193
460	<800	6800	6861	5942	5855
4973	4928	3983	3943	4168	4090
2725	2727	3427	3450	668	<800
4143	4099	4604	4590	4942	4844
1195	1183	4503	4403	15668	15466
2093	2068	5451	5369	2085	2079
5142	5207	22243	22372	388	<800
4103	4099	6052	6484	1093	1096
13196	12882	7746	7719	4098	4090
8147	8272	4644	4590	8408	8418
3525	3553	3988	3943	1344	1305
3631	3648	1909	1877	4105	4090
2235	2259	1095	1080	1263	1305
3215	3225	2019	2035	4483	4538
5186	5207	6503	6861	145	<800
30944	31100	7011	6861	392	<800
		652	<800	4589	4538
		663	<800	475	<800
		3627	3642	3456	3449
		815	800	860	854
		2140	2147	2511	2524
		2881	2829	5445	5390
		5472	5896	5792	5855
		786	<800	1343	1305

FEATURES

source
Location/Qualifiers
1. .176861
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/clone="RP11-532L16"
/clone_lib="RPC1 human BAC library 11"
BASE COUNT 53647 a 38409 c 35065 g 49740 t
ORIGIN

Query Match 68.0%; Score 17; DB 9; Length 176861;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 TGTGCCAGGACACGAA 17
|||||
Db 48753 TGTGCCAGGACACGAA 48769

RESULT 29
LOCUS

AL136090/c
AL136090/

DEFINITION

Human DNA sequence from clone RP11-97N19 on chromosome 20. Contains the 5' end of the SLC24A3 gene for solute carrier family 24 (sodium/potassium/calcium exchanger) member 3, ESTs, STSS, GSSs and a CpG island, complete sequence.

ACCESSION

AL136090

VERSION

AL136090.12 GI:9662903

KEYWORDS

HTG; CpG island; SLC24A3.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 185473)

AUTHORS

Williams, S.

JOURNAL

Direct Submission

Submitted (20-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

requests: clonerequest@sanger.ac.uk

On Aug 2, 2000 this sequence version replaced gi:9621474.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

only a small overlap as described above.

The following abbreviations are used to associate primary accession

numbers given in the feature table with their source databases:

En: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information

on the WORMPEP database can be found at

http://www.sanger.ac.uk/projects/c.elegans/wormpep This sequence

was generated from part of bacterial clone contigs of human

chromosome 20, constructed by the Sanger Centre Chromosome 20

Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr20

IMPORTANT: This sequence is not the entire insert of clone

RP11-97N19 It may be shorter because we sequence overlapping

sections only once, except for a 100 base overlap.

The true left end of clone RP11-97N19 is at 1 in this sequence. The

true right end of clone RP11-3K6 is at 52986 in this sequence.

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest. RP11-97N19 is from

the library RPI-11.1 constructed by the group of Pieter de Jong.
For further details see
<http://www.chori.org/bacpac/home.htm>

VECTOR: pBacE3.6.

FEATURES

```

source
1..185473
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/clone="RP11-97N19"
/clone.lib="RPI-11.1"
misc_feature
15..382
/note="match: STS: Em:G56094
match: GSS: Em:AQ319526"
399..699
/note="AluJo repeat: matches 1..299 of consensus"
1194..1509
/note="AluYb repeat: matches 1..310 of consensus"
complement(2925..3402)
/note="match: GSS: Em:AQ724001"
2961..3345
/note="THE1C repeat: matches 1..371 of consensus"
5053..5366
/note="AluJo repeat: matches 1..306 of consensus"
7059..7212
/note="AluJo repeat: matches 1..306 of consensus"
/note="L1ME repeat: matches 5518..5785 of consensus"
8415..8872
/note="match: GSS: Em:AQ463733"
8871..9370
/note="MER90 repeat: matches 35..563 of consensus"
9598..9731
/note="MIR repeat: matches 89..232 of consensus"
10480..10614
/note="MER5A repeat: matches 1..149 of consensus"
10620..10743
/note="AluY repeat: matches 1..129 of consensus"
10783..10973
/note="AluY repeat: matches 121..311 of consensus"
10984..11016
/note="MER5A repeat: matches 156..189 of consensus"
11857..12446
/note="L1MC/D repeat: matches 4381..4948 of consensus"
12452..12487
/note="19 copies 2 mer aa 100% conserved"
12514..12850
/note="L1MC/D repeat: matches 4978..5313 of consensus"
12889..13096
/note="L1MC/D repeat: matches 5320..5530 of consensus"
13104..13300
/note="L1PBa repeat: matches -1545..-1348 of consensus"
13309..13850
/note="L1PB1 repeat: matches -940..-402 of consensus"
13841..19303
/note="L1PB1 repeat: matches -265..6152 of consensus"
19323..19900
/note="L1MC2 repeat: matches 5525..6139 of consensus"
20697..21023
/note="THE1B repeat: matches 1..355 of consensus"
21024..22593
/note="THE1B-INTERNAL repeat: matches 3..1580 of
consensus"
22594..22972
/note="THE1B repeat: matches 1..364 of consensus"
23075..23384
/note="MER1B repeat: matches 2..331 of consensus"
23460..23804
/note="MLT1A1 repeat: matches 1..365 of consensus"
24745..25098
/note="MLT1A2 repeat: matches 3..374 of consensus"
25272..25696
/note="LTR33 repeat: matches 41..485 of consensus"
25697..25851
/note="MLT1A1 repeat: matches 1..155 of consensus"

```

```

repeat_region
25852..26158
/note="AluSc repeat: matches 1..309 of consensus"
26159..26367
/note="MLT1A1 repeat: matches 155..363 of consensus"
26368..26403
/note="LTR33 repeat: matches 8..41 of consensus"
26706..27162
/note="HERVL repeat: matches 4320..4790 of consensus"
27693..27870
/note="MIR repeat: matches 20..194 of consensus"
complement(29124..29524)
/note="match: GSS: Em:AQ736831"
29206..29277
/note="L2 repeat: matches 2470..2534 of consensus"
30160..30193
/note="HSMAR1 repeat: matches 1..34 of consensus"
30281..30354
/note="L2 repeat: matches 2671..2745 of consensus"
32566..32762
/note="L1ME3 repeat: matches 5881..6082 of consensus"
32764..32853
/note="L1ME3 repeat: matches 5615..5709 of consensus"
33032..33160
/note="MIR repeat: matches 42..169 of consensus"
33171..33294
/note="MER97a repeat: matches 767..880 of consensus"
33295..33584
/note="AluY repeat: matches 1..291 of consensus"
33585..34192
/note="MER97a repeat: matches 89..767 of consensus"
34283..35426
/note="L2 repeat: matches 1543..2744 of consensus"
35427..35779
/note="MLT1A1 repeat: matches 5..365 of consensus"
35780..36007
/note="L2 repeat: matches 1321..1543 of consensus"
36071..36127
/note="19 copies 3 mer tta 70% conserved"
37173..37267
/note="L2 repeat: matches 1693..1780 of consensus"
37268..37477
/note="MLT1A1 repeat: matches 158..365 of consensus"
37478..37781
/note="AluYa5 repeat: matches 1..301 of consensus"
37782..37938
/note="MLT1A1 repeat: matches 1..158 of consensus"
37939..38691
/note="L2 repeat: matches 1780..2592 of consensus"
38715..41361
/note="Charlie3 repeat: matches 1..2710 of consensus"
41401..41486
/note="MIR repeat: matches 81..167 of consensus"
41530..41689
/note="L2 repeat: matches 2590..2749 of consensus"
41860..42248
/note="L1PA7 repeat: matches 5756..6143 of consensus"
42261..42422
/note="FRAM repeat: matches 1..163 of consensus"
42780..43076
/note="AluY repeat: matches 1..299 of consensus"
43560..43875
/note="AluSq repeat: matches 1..307 of consensus"
44282..44374
/note="MIR repeat: matches 101..194 of consensus"
44573..45448
/note="L1PA2 repeat: matches 5262..6146 of consensus"
45448..46162
/note="L1PA2 repeat: matches 4528..5242 of consensus"
48401..48761
/note="THE1B repeat: matches 1..364 of consensus"
complement(48756..49120)
/note="match: GSS: Em:AQ017019"
49501..49791

```

```

/Note="Alusq repeat: matches 11. .297 of consensus"
49904. .50029
/Note="MER5B repeat: matches 47. .176 of consensus"
50069. .50360
/Note="Alusx repeat: matches 1. .293 of consensus"
51322. .51752
/Note="MER63B repeat: matches 20. .428 of consensus"

Query Match      68.0%; Score 17; DB 9; Length 185473;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CCAGGACCAAGAAATTT 21
|||||
Db 34147 CCAGGACCAAGAAATTT 34131

RESULT 30
HUAC004382/c 217873 bp DNA linear PRI 23-NOV-1999
LOCUS HUAC004382 Homo sapiens Chromosome 16 BAC clone C1987SK-A-152E5, complete
sequence.
ACCESSION AC004382
VERSION AC004382.1 GI:3252819
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Lofthus,B.J., Kim,U.J., Sneddon,V.P., Kalush,F., Brandon,R.,
Fuhrmann,J., Mason,T., Crosby,M.L., Barnstead,M., Cronin,L.,
Deslattes Mays,A., Cao,Y., Xu,R.X., Kang,H.L., Mitchell,S.,
Eichler,E., Harris,P.C., Venter,J.C. and Adams,M.D.
Genome duplications and other features in 12 Mb of DNA sequence
from human chromosome 16p and 16q
Genomics 60 (3), 295-308 (1999)
99425270
MEDLINE 10493829
PUBMED 2 (bases 1 to 217873)
ADAMS,M.D., Loftus,B.J., Zhou,L., Crosby,M., Fuhrmann,J.,
Mason,T.M., Brandon,R., Kim,U.J., Kerlavage,A.R. and Venter,J.C.
Homo sapiens Chromosome 16 BAC clone C1987SK-A-152E5
Unpublished
3 (bases 1 to 217873)
ADAMS,M.D. and Loftus,B.J.
Direct Submission
Submitted (07-MAR-1998) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, Email:
bjloftus@tigr.org
4 (bases 1 to 217873)
ADAMS,M.D.
Direct Submission
Submitted (24-JUN-1998) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
5 (bases 1 to 217873)
ADAMS,M.D.
Direct Submission
Submitted (25-JUN-1998) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
On Jun 24, 1998 this sequence version replaced gi:3172165.
Address all correspondence to: Mark Adams The Institute for Genomic
Research 9712 Medical Center Dr, Rockville, MD 20850, USA e-mail
address: mdadams@tigr.org. The orientation of the sequence is
from SP6 end to 3' end. Genes were identified by a combination of
five methods including: XGRAIL (available by anonymous ftp from
arthur.epm-ornl.gov), Genefinder (Phil Green, University of
Washington), Genscan (Chris Burge,
http://genome.stanford.edu/~chris/GENSCANW.html) searches of the
complete sequence against a peptide database, and the Human gene
Index database at TIGR (http://www.tigr.org/tdb/hgi/hgi.html).
Genes without peptide homology having spliced EST hits are termed
'Unknown gene product'. Genes encoding tRNAs are predicted by

```

```

TRNAScan-SE (Sean Eddy, http://genome.wustl.edu/eddy/TrNAScan-SE/).
Location/Qualifiers
1. .217873
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q13/21"
/clone="A-152E5"
8227. .16211
/gene="A-152E5.1"
Join(8827. .8918,10466. .10589,13523. .16211)
/gene="A-152E5.1"
Join(8846. .8918,10466. .10589,13523. .13607)
/gene="A-152E5.1"
/codon_start=1
/product="CC chemokine STCP-1"
/protein_id="AAC24306.1"
/db_xref="GI:3252820"
/translation="MARLOTALLVLLVLAVALQATEAGPYGANNEDSVCCRDYVRYR
LPLRVKHFVWTSDCSPFGVLLTFPRKELCADPRVFWKMLNKLKLSQ"
22519. .35075
/gene="A-152E5.2"
Join(22519. .22658,29654. .29774,32052. .35075)
/gene="A-152E5.2"
Join(22589. .22658,29654. .29774,32052. .33054)
/gene="A-152E5.2"
/codon_start=1
/product="CX3C chemokine precursor"
/protein_id="AAC24307.1"
/db_xref="GI:3252821"
/translation="KAPISLSMLRLATFOHLTVLLAGOHGVTKCNITCSKMTSKIP
VALLIYQOQASCGKRAITLLETRHRLFCADPKQWVKDAMOHLDROAAALTRNGT
FKQIGVKRPTTPAGGMDSVLEPEATGSSSLETPSSQARALGTSPFLPTG
VTSQSGRLPTPKAQDGPVTELFPRVPSVTAAYWQSAPHPGPGFSLWAKATISA
PSTQDPTQASTASSPAPENAPSEGORVWQGGQSPRENSLEREMGVPAPHTAQ
DWPGSAHVSVVVPVSESGTSPREPASGWTPKAEP IAHATMDPQRLGLVITPVPOA
QAATRRQAVGLLAFGLLFCLGVAFTYQSLQCPKRMAGMAGRLYIPRSCGSNSV
VLVPV"
63650. .65825
/gene="A-152E5.3"
Join(63650. .63771,64844. .64961,65528. .65825)
/gene="A-152E5.3"
Join(63702. .63771,64844. .64961,65528. .65624)
/gene="A-152E5.3"
/codon_start=1
/product="T cell-directed CC chemokine"
/protein_id="AAC24308.1"
/db_xref="GI:3252822"
/translation="MAPLKMLALVTLGLGASLQHIHAAGTINVGRECCLEYFKGAIPL
RKLAITYQTSQEDCSRDALVFTYVQGRAICSDPNKRVKNAVKYLQSLERS"
complement(77937. .>90706)
/gene="A-152E5.9"
complement(Join(77937. .79048,80026. .80107,80919. .81238,
82255. .82328,83818. .83986,86434. .86510,88960. .89073,
90550. .>90706))
/gene="A-152E5.9"
complement(Join(77937. .79048,80026. .80107,80919. .81034,
82255. .82328,83818. .83986,86434. .86510,88960. .89112,
90550. .>90706))
/gene="A-152E5.9"
complement(Join(78938. .79048,80026. .80107,80919. .81238,
82255. .82328,83818. .83986,86434. .86510,88960. .89073,
90550. .>90706))
/gene="A-152E5.9"
/codon_start=1
/product="Unknown gene product"
/protein_id="AAC24312.1"
/db_xref="GI:3252826"
/translation="MADFGISAGQFVAVVWDKSPVEALKGLVDKLOALTGNEGRYSV
ENIKQLLOCLVPGSTLTHSAEILAEIARTLRPGCLFLKEPVETADNNKSVKATSKL
CSALTSGLVEYKELOREPLTPEVOSVREHLGHESNLLFVQTKKPNFEVGSRRQ
LKLSITKKSPPSVKFAVDPAALAKLTLANDMEDDSMCFCCGSLTHRWPLEHYRLN
MNLNKEDRVDTFFTLDSKFLPACSHFSFLAETTTVSLIALNTLQDLIDSBELDDP

```

CDS
EDLKKPDPASIRAAACGGCKKCKKCKNCTGCTGAAIELEKEKSRQMSQPKSACGNCYL
GDAFCACSYLGMFAFAGKVKLLSDNLHDA"
complement(join(78938..79048,80026..80107,80919..81034,
82255..82328,83818..83986,86434..86510,88960..89112,
90550..90706))
/gene="A-152E5.9"
/codon_start=1
/product="Unknown gene product"
/protein_id="AAC24311.1"
/db_xref="GI:3252825"
/translation="MADEGISAGFVAVVMDKSSPVEALKGLVQKLOALITGNEGRSVV
ENIKQLQSAHKESFFDIILSLVPGSTVILHSABIELAEIARILRPGGCLFAKEPVETA
VDNNKVKYTKSLKLSALISLVEVKELELTPEEVQSVREHLGHSDNLILFVQIT
GKKNFEGSSQRLKLSITKSSPSVKPAVDPAARAKWTLSANDMEDDSMLIDSDLE
LDPELAKPDPASIRAAACGGCKKCKNCTGCTGAAIELEKEKSRQMSQPKSACGN
CYLDGAFRCASCPYLGMPAFKPGKVKLLSDNLHDA"
93011..93128
/note="7608, STS1-CSRL-24gl-1UA/CSRL-24gl-1uZ, Chr. -, Homo
sapiens"
/db_xref="dbSTS:G02122"
/db_xref="db_xref=.111048
/gene="A-152E5.11"
Join(<97284..97356,100817..100985,102578..102713,
106281..106423,106708..106792,108023..108127,
109342..109497,109822..109875,110360..111048)
/gene="A-152E5.11"
Join(97284..97356,100817..100985,102578..102713,
106281..106423,106708..106792,108023..108127,
109342..109497,109822..109875,110360..110395)
/gene="A-152E5.11"
/codon_start=1
/product="Unknown gene product"
/protein_id="AAC24313.1"
/db_xref="GI:3252827"
/translation="MAAAVSGALGRAGWLLQLRCLPVARCQALVPRAFHASAVGL
RSSDEQKQPPNSFSQOHSQAEKPPDESSHSPRYTDQGEEDIESEQLQHR
ILTALEFPVAGHTATAIAGASQSLGSSAASMGKDSILLHFVTCNRLTRY
LEEOKLVOLGAERKTDQFLDAVETRLRLPIYIEHWPRALSILMLPNIIPSSL
LLTSWDDMHWIAGDSTDFNWTYTRAMLAAYINTLEWMDQSDPDTWRFLENR
VNDAMNGHTAKQVKGSTGEALVQGLGAAVTLKMLTGLNQR"
112455..112519
/gene="A-152E5.7"
Join(112455..112581,112800..112849,115724..115792,
115946..115998,118934..119062,119479..119530,
119730..119898,120049..120123,120744..122519)
/gene="A-152E5.7"
Join(112496..112581,112800..112849,115724..115792,
115946..115998,118934..119062,119479..119530,
119730..119898,120049..120123,120744..120886)
/gene="A-152E5.7"
/codon_start=1
/product="RNA polymerase II subunit hRPB33"
/protein_id="AAC24309.1"
/db_xref="GI:3252823"
/translation="MPYANOPTVITELTDENVKFIENLDAVANSIRRVFAIEVPI
IALDWIDANSSVLHDEFAHGLIPLISDIVDKLOYSDCTCEFCPCSVET
LDVRNEDPHTVSRDLSSNPVPTVRNRNDPNPDYEQDILILKRGQELR
LRAAKGFGKEAKWNPAGVAFYDPDNLARHTVYKPEWPKSEYSELDSESOA
PDPNGKPERFYNVESCGSRPITIVLSALSLGKKLSDLQTLSEIQSDVLIN"
complement(122429..>129228)
/gene="A-152E5.8"
complement(join(122429..123270,123388..123511,
123676..123814,124568..124757,124875..124994,
125281..125395,125616..125723,129208..>1294228))
/gene="A-152E5.8"
complement(join(123152..123270,123388..123511,
123676..123814,124568..124757,124875..124994,
125281..125395,125616..125723,129208..1294228))
/gene="A-152E5.8"
/codon_start=1
/product="Unknown gene product"
/protein_id="AAC24310.1"
/db_xref="GI:3252824"

Query Match 58.0%; Score 17; DB 9; Length 217873;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 TGCAGGACACAGAAAT 19
|||||
Db 215636 TGCAGGACACAGAAAT 215620

RESULT 31
AC122476/c
LOCUS

DEFINITION Mus musculus chromosome UNK clone RP24-33412, WORKING DRAFT
SEQUENCE, 26 unordered pieces.

ACCESSION AC122476
VERSION AC122476.1 GI:21105937
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Mus musculus.
ORGANISM Mus musculus.

REFERENCE

AUTHORS 1 (bases 1 to 220388)
TITLE McPherson,J.D. and Waterston,R.H.
JOURNAL The sequence of Mus musculus clone

REFERENCE 2 (bases 1 to 220388)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission

JOURNAL Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

COMMENT

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: M_BB0334102
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 210112 bases at least Q40
Consensus quality: 212638 bases at least Q30
Consensus quality: 214538 bases at least Q20
Insert size: 197000; agarose-fp
Insert size: 220179; sum-of-contigs
Quality coverage: 8.38 in Q20 bases; agarose-fp
Quality coverage: 6.32 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 26 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

* 1 1209: contig of 1209 bp in length
* 1210: gap of unknown length
* 1310: contig of 1158 bp in length
* 2468: gap of unknown length
* 2568: contig of 1000 bp in length
* 3568: gap of unknown length
* 3668: contig of 1173 bp in length
* 4841: gap of unknown length
* 6056: contig of 1116 bp in length
* 6157: gap of unknown length
* 7607: contig of 1451 bp in length

```

* 7608 7707: gap of unknown length
* 7708 8868: contig of 1161 bp in length
* 8869 8968: gap of unknown length
* 8969 10168: contig of 1200 bp in length
* 10169 10268: gap of unknown length
* 10269 11426: contig of 1158 bp in length
* 11427 11526: gap of unknown length
* 11527 12871: contig of 1345 bp in length
* 12872 12971: gap of unknown length
* 12972 14451: contig of 1480 bp in length
* 14452 14551: gap of unknown length
* 14552 16718: contig of 2167 bp in length
* 16719 16818: gap of unknown length
* 16819 18827: contig of 2009 bp in length
* 18828 18927: gap of unknown length
* 18928 21908: contig of 2981 bp in length
* 21909 22008: gap of unknown length
* 22009 26312: contig of 4304 bp in length
* 26313 26412: gap of unknown length
* 26413 29890: contig of 3478 bp in length
* 29891 36126: contig of 6136 bp in length
* 36127 36226: gap of unknown length
* 36227 41565: contig of 5339 bp in length
* 41566 41665: gap of unknown length
* 41666 49129: contig of 7464 bp in length
* 49130 49229: gap of unknown length
* 49230 66134: contig of 16905 bp in length
* 66135 66234: gap of unknown length
* 66235 81647: contig of 15413 bp in length
* 81648 88332: contig of 16585 bp in length
* 88333 98432: gap of unknown length
* 98433 120029: contig of 21597 bp in length
* 120030 120129: gap of unknown length
* 120130 147463: contig of 27334 bp in length
* 147464 147563: gap of unknown length
* 147564 180750: contig of 33187 bp in length
* 180751 180850: gap of unknown length
* 180851 220388: contig of 39538 bp in length.
FEATURES
Source
Location/Qualifiers
1. .220388
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP24-334L2"
misc_feature 1..1209
/note="assembly_name:Contig30"
misc_feature 1310..2467
/note="assembly_name:Contig33"
misc_feature 2568..3567
/note="assembly_name:Contig40"
misc_feature 3668..4840
/note="assembly_name:Contig45"
misc_feature 4941..6056
/note="assembly_name:Contig55"
misc_feature 6157..7607
/note="assembly_name:Contig62"
misc_feature 7708..8868
/note="assembly_name:Contig63"
misc_feature 8969..10168
/note="assembly_name:Contig66"
misc_feature 10269..11426
/note="assembly_name:Contig67"
misc_feature 11527..12871
/note="assembly_name:Contig69"
misc_feature 12972..14451
/note="assembly_name:Contig70"
misc_feature 14552..16718
/note="assembly_name:Contig71"
misc_feature 16819..18827
/note="assembly_name:Contig72"
18928..21908

```

```

/note="assembly_name:Contig73"
22009..26312
/note="assembly_name:Contig74"
26413..29890
/note="assembly_name:Contig75"
2991..36126
/note="assembly_name:Contig76"
36227..41565
/note="assembly_name:Contig77"
41666..49129
/note="assembly_name:Contig78"
49230..66134
/note="assembly_name:Contig79"
66235..81647
/note="assembly_name:Contig80"
clone_end:SP6
vector_side:right"
81748..98332
/note="assembly_name:Contig81"
98433..120029
/note="assembly_name:Contig82"
120130..147463
/note="assembly_name:Contig83"
147564..180750
/note="assembly_name:Contig84"
180851..220388
/note="assembly_name:Contig85"
clone_end:rf7
vector_side:left"
BASE COUNT 65307 a 41282 c 42284 g 68931 t 2584 others
ORIGIN
Query Match 68.0% Score 17: DB 2; Length 220388;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 GCCAGGAACACGAAATT 20
|||||
Db 130456 GCCAGGAACACGAAATT 130440

```

```

RESULT 32
AC122295/c
LOCUS AC122295 231923 bp DNA linear HTG 25-MAY-2002
DEFINITION Mus musculus chromosome UNK clone RP23-263E3, WORKING DRAFT
SEQUENCE, 7 unordered pieces.
ACCESSION AC122295
VERSION AC122295.2 GI:21206454
KEYWORDS HTG; HIGS-PHASE1; HTGS-DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 231923)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE The sequence of Mus musculus clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 231923)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 3 (bases 1 to 231923)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (25-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
COMMENT On May 25, 2002 this sequence version replaced gi:21105151.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml

```

Contact: submissions@watson.wustl.edu

 Project information
 Center project name: M_BA0263E03

 Summary Statistics

 Sequencing vector: M13; 0%
 Chemistry: Dye-primer ET; 100%
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 228616 bases at least Q40
 Consensus quality: 229657 bases at least Q30
 Consensus quality: 230256 bases at least Q20
 Insert size: 18200; agarose-fp
 Insert size: 231323; sum-of-contigs
 Quality coverage: 15.81 in Q20 bases; agarose-fp
 Quality coverage: 11.58 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 7 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 1096: contig of 1096 bp in length
 * 1097 1196: gap of unknown length
 * 1197 4371: contig of 3175 bp in length
 * 4372 4471: gap of unknown length
 * 4472 26809: contig of 22338 bp in length
 * 26810 26909: gap of unknown length
 * 26910 54548: contig of 27639 bp in length
 * 54549 54648: gap of unknown length
 * 54649 91374: contig of 36726 bp in length
 * 91375 91474: gap of unknown length
 * 91475 141654: contig of 50180 bp in length
 * 141655 141754: gap of unknown length
 * 141755 231923: contig of 90169 bp in length.
 *
 * Location/Qualifiers
 * 1. 231923
 * /organism="Mus musculus"
 * /db_xref="taxon:10090"
 * /chromosome="UNK"
 * /clone="RP23-263E3"
 *
 * misc_feature
 * 1. .1096
 * /note="assembly_name:Contig11"
 * 1197. .4371
 * /note="assembly_name:Contig12"
 * 4472. 26809
 * /note="assembly_name:Contig13"
 * 26910. .54548
 * /note="assembly_name:Contig14"
 * 54649. .91374
 * /note="assembly_name:Contig15"
 * 91475. 141654
 * /note="assembly_name:Contig16"
 * 141755. 231923
 * /note="assembly_name:Contig17"
 * 70704 a 44374 c 44153 g 72072 t 620 others
 *
 * BASE COUNT 70704 a 44374 c 44153 g 72072 t
 * ORIGIN
 *
 * Query Match 68.0%; Score 17; DB 2; Length 231923;
 * Best Local Similarity 100.0%; Pred. No. 9.7;
 * Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 *
 * QY 5 CCAGGACCAAGAAATTT 21
 * |||||
 * Db 167119 CCAGGACCAAGAAATTT 167103
 *
 * RESULT 33
 * CEY111B2A/c

LOCUS CEY111B2A 257703 bp DNA linear INV 13-AUG-2002
 DEFINITION Caenorhabditis elegans cosmid Y111B2A, complete sequence.
 ACCESSION ALI32904 ALI32885 ALI32886 ALI32889 ALI32890 ALI32906 ALI32908
 ALI32945
 VERSION ALI32904.4 GI:21615504
 KEYWORDS HTG.
 SOURCE Caenorhabditis elegans.
 ORGANISM Caenorhabditis elegans.
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida;
 Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
 1 none.
 REFERENCE Genome sequence of the nematode C. elegans: a platform for
 AUTHORS investigating biology. The C. elegans Sequencing Consortium
 TITLE Science 282 (5396), 2012-2018 (1998)
 JOURNAL 99069613
 MEDLINE The C.elegans Sequencing Consortium.
 REMARK 2 (bases 1 to 257703)
 REFERENCE Sulston, J.E.
 AUTHORS Direct Submission
 TITLE Submitted (13-AUG-2002) Nematode Sequencing Project, Sanger
 JOURNAL Institute, Hinxton, Cambridge CB10 1SA, England and Department of
 Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:
 jes@sanger.ac.uk or rvnematode.wustl.edu
 On Jun 26, 2002 this sequence version replaced gi:13548453.
 COMMENT Coding sequences below are predicted from computer analysis, using
 predictions from GeneFINDER (P. Green, U. Washington), and other
 available information.
 Current sequence finishing criteria for the C. elegans genome
 sequencing consortium are that all bases are either sequenced
 unambiguously on both strands, or on a single strand with both a
 dye primer and dye terminator reaction, from distinct subclones.
 Exceptions are indicated by an explicit note.
 For a graphical representation of this sequence and its analysis
 see: - http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?
 name=Y111B2A
 IMPORTANT: This sequence is NOT necessarily the entire insert of
 the specified clone. It may be shorter because we only sequence
 overlapping sections once, or longer because we arrange for a small
 overlap between neighbouring submissions.
 IMPORTANT: This sequence is not the entire insert of clone Y111B2A.
 It may be shorter because we only sequence overlapping sections
 once, or longer because we only sequence overlapping sections
 neighbouring submissions.
 The true right end of clone Y49E10 is at 105 in this sequence. The
 start of this sequence (1. .105) overlaps with the end of sequence
 Z98866.
 The end of this sequence (257598. .257703) overlaps with the start
 of sequence Z93397.
 FEATURES
 Location/Qualifiers
 1. 257703
 /organism="Caenorhabditis elegans"
 /db_xref="taxon:6239"
 /chromosome="III"
 /clone="Y111B2A"
 complement(join(35. .160,217. .298,1625. .2304,2347. .2427,
 2477. .2541,2602. .2749))
 /gene="Y111B2A.1"
 complement(join(35. .160,217. .298,1625. .2304,2347. .2427,
 2477. .2541,2602. .2749))
 /gene="Y111B2A.1"
 /note="contains similarity to Pfam domain: PF00069
 (Eukaryotic protein kinase domain), Score=131.2,
 E-value=6.2e-36, N=1"
 /codon_start=1
 /protein_id="CAC35832.1"
 /db_xref="GI:13548455"
 /db_xref="SPTREMBL:Q9BHM0"
 /translation="MKRSTCLOHSLPAKKAYSSCGEKLFRFVDPHGTIDKSRFT
 NWGTDLENVSLGEGVGFVFEVADAEKKKFKALKFKAGHTQFCEDWVRESLSIP
 NYLQIMKSVAAAHHPNLIKLVSGKMAVPEGYRIAILMELRGKSLFSYMDTKSEV
 PKFEISFPIATIREIGAOILSKMKLEKMEIVHLDKPENICFTSTCTFTEIRKDYC
 YISPSLELVCIIDFGNARKSQEDKSSVVTQNTQYRAPEIFMGLPFSTRSDVNSFG"

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Lamiales; Orobanchaceae; Orobanchae. 1 (bases 1 to 286)

REFERENCE
AUTHORS
TITLE
JOURNAL

Direct Submission
Submitted (07-AUG-1998) Lohan A.J., Smurfit Institute of Genetics, Trinity College Dublin, University of Dublin, Trinity College, Dublin 2, IRELAND

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

2 (bases 1 to 286)
Lohan A.J. and Wolfe K.H.
A subset of conserved tRNA genes in plastid DNA of nongreen plants
Genetics 150 (1), 425-433 (1998)

98393598
9725858

FEATURES
source

Location/Qualifiers

1..286
/organism="Orobanchae minor"
/organelle="plastid:chloroplast"
/db_xref="taxon:36748"
27..99
/gene="trnF(GAA)"
27..99
/gene="trnF(GAA)"
/product="tRNA-Phe"
/note="codon recognized: UUC"
/anticodon="(pos:60..62,aa:Phe)"
94 a 36 c 52 g 104 t

Query Match

Best Local Similarity 64.0%; Score 16; DB 8; Length 286;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGACCCAGA 16
|||||

Db 101 TGTGCCAGGACCCAGA 86

RESULT 35
CH2M05/C

LOCUS

Chloroplast Zea mays gene encoding tRNA specific for Phenylalanine.
(tRNA-Phe).
320 bp DNA linear PLN 31-MAR-1992

DEFINITION

V00174

ACCESSION

V00174.1 GI:12397

VERSION

transfer RNA; transfer RNA-Phe.

KEYWORDS

Ze mays.

SOURCE

ORGANISM

Chloroplast Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 320)

Steinmetz, A., Krebbers, E.T., Schwarz, Z., Gubbins, E.J. and Bogorad, L.

Nucleotide sequences of five maize chloroplast transfer RNA genes and their flanking regions

J. Biol. Chem. 258 (9), 5503-5511 (1983)

83213507

6853530

FEATURES

source

Location/Qualifiers

1..320
/organism="Zea mays"
/organelle="plastid:chloroplast"
/db_xref="taxon:4577"
81..153
/product="tRNA-Phe"
114 a 43 c 60 g 103 t

trNA

BASE COUNT 114 a 43 c 60 g 103 t

ORIGIN

Query Match 64.0%; Score 16; DB 8; Length 320;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGACCCAGA 16
|||||

Db 155 TGTGCCAGGACCCAGA 140

RESULT 36

AX017232/c

LOCUS

Sequence 16 from Patent WO947670.

DEFINITION

AX017232

ACCESSION

AX017232.1 GI:10042152

VERSION

human.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 585)

Tate, S.N., Grose, D.T. and Hick, C.A.

Mammalian sodium channel proteins

Patent: WO 947670-A 16 23-SEP-1999;

TATE SIMON NICHOLAS (GB); GLAXO GROUP LTD (GB); GROSE DAVID THOMAS (GB); HICK CAROLINE ANNE (GB)

FEATURES

source

Location/Qualifiers

1..585
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 181 a 117 c 112 g 175 t

ORIGIN

Query Match 64.0%; Score 16; DB 6; Length 585;

Best Local Similarity 100.0%; Pred. No. 59;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 GGAACCCAGAAATTTAC 23

|||||

Db 471 GGAACCCAGAAATTTAC 456

RESULT 37

SOTRNAPLE/c

LOCUS

S.oleracea chloroplast DNA for phe-tRNA and leu-tRNA.

DEFINITION

X87343

ACCESSION

X87343.1 GI:840898

VERSION

transfer RNA-Leu; transfer RNA-Phe; tRNA-Leu gene; tRNA-Phe gene.

KEYWORDS

SOURCE

ORGANISM

Chloroplast Spinacia oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.

1 (bases 1 to 1147)

Lin, C.H., Wu, C.Y. and Chen, L.J.

Characterization of tRNAs maturation of spinach chloroplast

Unpublished

2 (bases 1 to 1147)

Lin, C.H. and Wu, C.Y.

Direct Submission

Submitted (22-MAY-1995) C.H. Lin, Institute of Molecular Biology, National Chung Hsing University, Taichung, 40227 Taiwan, R.O.C

FEATURES

source

Location/Qualifiers

1..1147
/organism="Spinacia oleracea"
/organelle="plastid:chloroplast"
/db_xref="taxon:3562"
102..151
/gene="tRNA-Leu"
102..151
/gene="tRNA-Leu"
/product="tRNA-Leu"
/note="3' exon"
488..561
/gene="tRNA-Phe"

gene

trNA

BASE COUNT 114 a 43 c 60 g 103 t

ORIGIN

Query Match 64.0%; Score 16; DB 8; Length 320;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

trna 488..561
/gene="trna-Phe"
/product="trna-Phe"
BASE COUNT 397 a 172 c 184 g 394 t
ORIGIN

Query Match 64.0%; Score 16; DB 8; Length 1147;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGAACCCAGA 16
|||||
DB 562 TGTGCCAGAACCCAGA 547

RESULT 38
TOBCPLB3/c
LOCUS Tobacco Leu-trna and Phe-trna genes. 1268 bp DNA linear PLN 26-APR-1995
ACCESSION M16898
VERSION 1
KEYWORDS transfer RNA-Leu; transfer RNA-Phe.
SEGMENT 3 of 3
SOURCE Chloroplast Nicotiana tabacum (variety Bright Yellow 4) DNA.
ORGANISM

REFERENCE 1 (bases 1 to 1268)
Yamada,K., Shinozaki,K. and Sugura,M.
DNA sequences of tobacco chloroplast genes for trna-Ser (GGA),
trna-Thr (UGU), trna-Leu (UAA), trna-Phe (GAA): the trna-Leu gene
contains a 503 bp intron
PLant Mol. Biol. 6, 193-199 (1986)
JOURNAL The Leu-trna anticodon (TAA) begins with the last bp (119) of exon
COMMENT 1, and ends with the first two bp (623-624) of exon 2.
FEATURES
Location/Qualifiers
1..1268
/organism="Nicotiana tabacum"
/organelle="plastid:chloroplast"
/variety="Bright Yellow 4"
/db_xref="taxon:4097"
join(85..119,623..673)
/gene="Leu-trna"
join(85..119,623..673)
/product="trna-Leu"
85..119
/gene="Leu-trna"
/note="putative; does not fit consensus"
/cons_splice=(5'site:no, 3'site:no)
623..673
/gene="Leu-trna"
/note="putative"
/number=2
1029..1101
/gene="Phe-trna"
1029..1101
/product="trna-Phe"
/note="codon recognized: UUC"
anticodon=(pos:1062..1064,aa:Phe)
223 c 226 g 376 t
BASE COUNT 443 a 223 c 226 g 376 t
ORIGIN About 0.7 kb after segment 2.

Query Match 64.0%; Score 16; DB 8; Length 1268;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGAACCCAGA 16
|||||
DB 562 TGTGCCAGAACCCAGA 547

Query Match 64.0%; Score 16; DB 8; Length 1268;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGAACCCAGA 16
|||||
DB 1103 TGTGCCAGAACCCAGA 1088

RESULT 39
PVTRPHE/c
LOCUS P.vulgaris trna-Phe gene and trna-Leu gene-3'end. 1358 bp DNA linear PLN 08-MAY-1992
ACCESSION X58537
VERSION 1
KEYWORDS transfer RNA-Leu; transfer RNA-Phe.
SOURCE Phaseolus vulgaris.
ORGANISM Chloroplast Phaseolus vulgaris
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Phaseolus.
1 (bases 1 to 1358)
Mubumbila,M.
Direct Submission
Submitted (13-MAR-1991) M. Mubumbila, Biochemistry Department,
University of Zimbabwe, PO BOX MP, 167 Mount Pleasant, Harare,
Zimbabwe
2 (bases 1 to 1358)
Mubumbila,M. and Mubumbila,M.
DNA heterogeneity in common bean chloroplasts
Unpublished
COMMENT for overlapping sequences see: X02444.
FEATURES
Location/Qualifiers
1..1358
/organism="Phaseolus vulgaris"
/organelle="plastid:chloroplast"
/strain="Saxa"
/db_xref="taxon:3885"
/tissue_type="leaves"
1..607
/number=1
608..657
/gene="trna-Leu"
608..657
/partial
/gene="trna-Leu"
608..657
/number=2
1015..1020
-35_signal
1039..1044
-10_signal
1077..1140
gene
/gene="trna-Phe"
1077..1140
trna
/gene="trna-Phe"
1141..1358
misc_feature
/note="intergenic spacer"
BASE COUNT 478 a 199 c 202 g 479 t
ORIGIN

Query Match 64.0%; Score 16; DB 8; Length 1358;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGAACCCAGA 16
|||||
DB 1142 TGTGCCAGAACCCAGA 1127

RESULT 40
MITATRSF/c
LOCUS MITATRSF 1608 bp DNA linear PLN 31-MAR-1992
DEFINITION Wheat mitochondrial DNA for transfer RNA-Ser(GGA) and -Phe(GAA).
ACCESSION X15118
VERSION X15118.1 GI:13735
KEYWORDS pseudogene; transfer RNA; transfer RNA-Phe; transfer RNA-Ser; trnf

gene; trns gene.
Triticum aestivum.
Mitochondrion Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Poales; Poaceae; Triticeae; Triticeae; Triticeae.
1 (bases 1 to 1608)
Gray M.W.
Direct Submission
Submitted (26-APR-1989) Gray M.W., Dalhousie University, Department
of Biochemistry, Halifax Nova Scotia B3H 4H7, Canada
2 (bases 1 to 1608)
Joyce, P.B. and Gray, M.W.
Chloroplast-like transfer RNA genes expressed in wheat mitochondria
Nucleic Acids Res. 17 (14), 5461-5476 (1989)
89345142
PUBMED
2762145
COMMENT
Data kindly reviewed (31-aug-1989) by Gray M.W.
FEATURES
source
1. .1608
Location/Qualifiers
/organism="Triticum aestivum"
/organelle="mitochondrion"
/db_xref="taxon:4565"
968..1054
/product="trna-Ser"
/note="codon recognized: GGA"
1107..1155
/note="trna-Leu pseudogene 3'-exon"
1470..1542
/product="trna-Phe"
/note="codon recognized: GAA"
522 a 304 c 281 g 501 t
BASE COUNT
ORIGIN
Query Match 64.0%; Score 16; DB 8; Length 1608;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGTGCCAGGACACCA 16
|||||
Db 1544 TGTGCCAGGACACCA 1529
RESULT 41
AF232699/c
LOCUS
AF232699 1700 bp mRNA linear MAM 21-FEB-2002
DEFINITION
Sus scrofa 3-beta-hydroxysteroid dehydrogenase/delta-5-delta-4
isomerase (3b-HSD) mRNA, 3b-HSD-1 allele, complete cds.
ACCESSION
AF232699
VERSION
AF232699.2 GI:18846439
SOURCE
Sus scrofa.
ORGANISM
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 1700)
von Teichman A., Joerg, H., Werner, P., Brenig, B. and Stranzinger, G.
cDNA cloning and physical mapping of porcine 3 beta-hydroxysteroid
dehydrogenase/Delta 5-Delta 4 isomerase
Anim. Genet. 32 (5), 298-302 (2001)
21540582
PUBMED
11683717
REFERENCE
2 (bases 1 to 1700)
von Teichman A.F., Joerg, H.W., Werner, P. and Stranzinger, G.F.
Direct Submission
Submitted (08-FEB-2000) Department of Animal Science, Swiss Federal
Institute of Technology Zurich, Tannenstrasse 1, Zurich 8092,
Switzerland
3 (bases 1 to 1700)
von Teichman A.F., Joerg, H.W., Werner, P. and Stranzinger, G.F.
Direct Submission
Submitted (21-FEB-2002) Department of Animal Science, Swiss Federal
Institute of Technology Zurich, Tannenstrasse 1, Zurich 8092,

Switzerland
Sequence update by submitter
On Feb 21, 2002 this sequence version replaced gi:7141331.
Location/Qualifiers
1. .1700
/organism="Sus scrofa"
/db_xref="taxon:9823"
/chromosome="4"
/map="4q16-q21"
/tissue_type="adipose"
1. .1700
/gene="3b-HSD"
/allele="1"
173..1294
/gene="3b-HSD"
/codon_start=1
/product="3-beta-hydroxysteroid
dehydrogenase/delta-5-delta-4 isomerase"
/protein_id="AAF37295.2"
/db_xref="GI:18846440"
/translation="MAGWSCIVTGGGGFLGQRIVHLLLEKDLQEIIRVLDKVPKPEVR
EFSKLSKIKLTMLEGDILDEQCLKAGCOGASVHTASIIDVNAVGRVTKVNV
KGTQLLEACVQASVPVFLHTSSIEVAGPNSREVTONACEDELETANSAPYLSKK
LAERAVLEANGWALONGTTLHTCALRPMWLYCGSPFIFAHMKALENGVLTNPKF
SRVNPVYGVNAWHILALRALRDPKALSVOGQFYVADDTPPQSDLDLNTTGKWK
GFLDSRSLPSPSLRWLAFLLEIVSELSPIYQPPFNEHFVTLCSVFTVSYKKA
QRDLGYEPLFTWEAAKQKTKAWGSLVKQKHEALKTKTH"
BASE COUNT 376 a 497 c 460 g 367 t
ORIGIN
Query Match 64.0%; Score 16; DB 4; Length 1700;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 TGCAGGACACCAAAA 18
|||||
Db 145 TGCAGGACACCAAAA 130
RESULT 42
RICMT16/c
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) mitochondrial gene for
tRNAs, complete sequences.
ACCESSION
D13112
VERSION
D13112.1 GI:287392
KEYWORDS
trna-leu; trna-phe; trna-Ser.
SOURCE
Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare)
mitochondrion DNA.
ORGANISM
Mitochondrion Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaeae; Oryza.
1
Nakazono, M. and Hirai, A.
REFERENCE
AUTHORS
TITLE
Identification of the entire set of transferred chloroplast DNA
sequences in the mitochondrial genome of rice
Mol. Gen. Genet. 236 (2-3), 341-346 (1993)
JOURNAL
MEDLINE
93173110
REFERENCE
2 (bases 1 to 4894)
Nakazono, M.
AUTHORS
JOURNAL
COMMENT
Submitted (28-Aug-1992) to DBJ by:
Atsushi Hirai
Laboratory of Radiation Genetics
Faculty of Agriculture,
The University of Tokyo
1-1-1 Yayoi Bunkyo-ku
Tokyo 113
Japan
Phone: 03-3812-2111
Fax: 03-3815-9335.
Location/Qualifiers

source 1. 4894
/organism="Oryza sativa (japonica cultivar-group)"
/organism="Mitochondrion"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
641. 4727
/note="homologous to ctDNA"
1145. 1231
/product="trna-Ser"
1458. 2058
/gene="rps4"
1458. 2058
/gene="rps4"
2404. 2458
/product="trna-Leu"
2769. 2841
/product="trna-Phe"
3398. 3885
/product="ORF159"
3983. 4717
/gene="ndhK"
3983. 4717
/gene="ndhK"
BASE COUNT 1523 a 941 c 864 g 1566 t
ORIGIN
Query Match 64.0%; Score 16; DB 8; Length 4894;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGTGCCAGGACACCA 16
|||||
Db 2843 TGTGCCAGGACCA 2828
RESULT 43
AF072704
LOCUS
DEFINITION
Zea mays stock Acl069 al::rdt sh2 Y21 (y21) gene, complete cds; and
disruption of al gene by rdt transposon.
ACCESSION
AF072704
VERSION
AF072704.1 GI:20385027
KEYWORDS
Zea mays.
SOURCE
Zea mays.
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 12558)
Yao, H., Zhou, Q., Li, J., Smith, H., Yandea, M., Nikolau, B.J. and
Schnable, P.S.
Molecular characterization of meiotic recombination across the
140-kb multigenic al-sh2 interval of maize
Proc. Natl. Acad. Sci. U.S.A. 99 (9), 6157-6162 (2002)
21980573
PUBMED
11959509
REFERENCE
2 (bases 1 to 12558)
Zhou, Q., Yao, H., Nikolau, B.J. and Schnable, P.S.
AUTHORS
Direct Submission
TITLE
Submitted (20-JUN-1998) Zoology and Genetics, Iowa State
University, Ames, IA 50011, USA
JOURNAL
Location/Qualifiers
FEATURES
1. 12558
/organism="Zea mays"
/db_xref="taxon:4577"
/chromosome="3"
/map="3L; between al and sh2"
/note="al::rdt sh2 stock; obtained from a 10-kb genomic
clone (pE10) derived from the al::rdt sh2 stock (XU et al.
1995). A 4.3-kb EcoRI/HindIII fragment of pE10 was
subcloned into pMOB (Gold Biotechnology, Inc., St. Louis,
MO) to create pEH4.3. Similarly, a 1.5-kb HindIII/SacI
fragment of pE10 was subcloned into pMOB to generate

gene 484. 3589
/gene="y21"
Join(484. 634,775. 944,1142. 1228,1774. 1935,2043. 2125,
2803. 2991,3068. 3589)
/gene="y21"
/product="y21"
Join(484. 634,775. 944,1142. 1228,1774. 1935,2043. 2125,
2803. 2991,3068. 3446)
/gene="y21"
/note="translation start position is predicted based on
conservation among rice, sorghum and maize"
/codon_start=1
/evidence=not_experimental
/product="Y21"
/protein_id="AA021160.1"
/db_xref="GI:20385028"
/translation="MAREAAEATAGSAGAGCEWSEALKSFLDHTPVSSYPGAL
OPTASPEVEKLVGSLGALDAMYSNAGAVITDVHSSLGKYVDRIQGVFFSSLV
LWALEEGKVESESDTDFLTKHQHQAETKIAWLAKLFWEFFPVRHDTLPH
AMLLFSKHLNVPVYVESMSSAIGFVTDAMVLELLQSGLEWLDKADKQISEFR
FANRKEVLVSDQTLADGLHLSKEKTGAVIDRKTSRLIGSLQSDLYFLDDSTL
FSKRTTLELISLNKTKSTAESNAPGRNIALALRSQSGMVGIPATNLESDT
LKQAMEKLALARSQSGSFVDEHGRVQGVVTRDIISVFPSPCSDRIDGATTFSAALE
QAGCRVENGOMFQN"
order(1319. 1321,1682. 1684)
/note="transposon MPF target site duplication"
/rpt_type=direct
1322. 1681
/note="miniature inverted-repeat transposable element"
/transposon="mpf"
order(1322. 1336,1667. 1681)
/note="mpf terminal inverted repeat"
/rpt_type=inverted
complement(3870. 3928)
/product="trna-Phe"
/evidence=not_experimental
9168. 10252
/gene="al"
/note="nonfunctional NADPH-dependent reductase; disrupted
due to transposon rdt insertion in exon 4"
allele="al::rdt"
9168. 9390
/gene="al"
/number=1
9486. 9655
/gene="al"
/number=2
9772. 9966
/gene="al"
/number=3
10113. 10252
/gene="al"
/note="disrupted due to transposon rdt insertion"
order(10245. 10252,10957. 10964)
/note="transposon rdt target site duplication"
/rpt_type=direct
10253. 10956
/transposon="rdt"
10957. 11628
/gene="al"
/note="nonfunctional NADPH-dependent reductase; disrupted
due to transposon rdt insertion in exon 4"
allele="al::rdt"
10957. 11628
repeat_region
repeat_region
repeat_region
trna
gene
exon
exon
exon
exon
exon
repeat_region
repeat_region
gene
exon

```

/feature="al"
/feature="disrupted due to transposon rdt insertion"
/number=4
BASE COUNT 3586 a 2608 c 2755 g 3609 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 47;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGACCAGA 16
|||||
Db 3868 TGTGCCAGGACCAGA 3883

RESULT 44
AF434192
LOCUS
DEFINITION
Zea mays line LH82 transposon Ins2, YZ1 (yz1) gene, YZ1-LH82
allele, complete cds; tRNA-Phe (trnF) gene, complete sequence;
retrotransposon Machiavelli gag and pol (gag/pol) gene, complete
cds; and retrotransposon-like Ozymandias and MITE Guatl, complete
sequence.
ACCESSION
AF434192
VERSION
AF434192.1 GI:20502801
KEYWORDS
Zea mays.
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 15783)
AUTHORS
Yao H., Zhou Q., Li J., Smith, H., Yandean, M., Nikolau, B.J. and
Schnable, P.S.
TITLE
Molecular characterization of meiotic recombination across the
140-kb multigenic al-sh2 interval of maize
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 99 (9), 6157-6162 (2002)
MEDLINE
21980573
PUBMED
11959909
REFERENCE
2 (bases 1 to 15783)
Yao H., Zhang, Y., Nikolau, B.J. and Schnable, P.S.
Direct Submission
Submitted (15-OCT-2001) Zoology and Genetics, Iowa State
University, Ames, IA 50011, USA
LOCATION/Qualifiers
1. 15783
/organism="Zea mays"
/isolate="line LH82"
/db_xref="taxon:4577"
/chromosome="3"
/map="between al and sh2"
/notes="obtained from the approximately 28-kb cosmid clone
Cos9-10 derived from YAC ASH-2 (Civardi et al., 1994). A
7.5-kb NotI/HindIII fragment from Cos9-10 that has about a
1.4-kb of overlap with the distal end of the al-mum2
sequence (GenBank Accession Number AF347696) was subcloned
into pMOB (Gold Biotechnology, Inc., St. Louis, MO) to
create phy53. A 4.7-kb HindIII fragment from Cos9-10 was
also subcloned into pMOB, generating phy62. A 2.7-kb
EcoRV fragment of Cos9-10 that overlaps both phy62 and
phy53 was subcloned into pMOB to generate phyEV.
Similarly, a 3.8-kb EcoRI fragment of Cos9-10 that
overlaps phy62 was subcloned into pCOS8-72 (GenBank
Accession Number AF061788) vector to generate PCS-YZ.
Clones phy53, phy62, phyEV, and PCS-YZ were used as
templates for sequencing using a transposon-mediated
sequencing procedure (STRATHMANN et al. 1991). In all
cases both strands were sequenced."
1021..1027
repeat_region
/notes="target site duplication"
/rpt_type=direct
1028..1395
repeat_region
/transposon="Ins2"

```

repeat_region

/note="target site duplication"

/rpt_type=direct

1649..4390

/gene="yz1"

/allele="YZ1-LH82"

join(1649..1799,1940..2109,2313..2402,2584..2745,

2853..2935,3604..3792,3869..4390)

/gene="yz1"

/product="YZ1"

join(1649..1799,1940..2109,2313..2402,2584..2745,

2853..2935,3604..3792,3869..4247)

/gene="yz1"

/codon_start=1

/product="YZ1"

/protein_id="AAM22634.1"

/db_xref="GI:20502802"

/translation="MAREAAEATASRESGAAGRCWEALKFLDHPVSPGAL

OPTASPAVEVKGVLGALDAMYSNAAGAVIIDVHSSLGKVDRIQGVFESLI

LWALELQVSEPTSDTSLTIKHOIAETKIAWLAKLFWPPFVPTHILF

HAWLFSKHHLNVAVPVSMSSNVIQVQDAMVLELLQSSGLEWLDKIDKQJSEF

RFANVRKPLVYSDTLDGLHILSKERTGVAVIDRTSRILGSLQCSLIYFLDDST

LFSKTTTLEELISLNKTRKSTAEACAPRNILALRSQSGMVGVPATNLESD

TLQQAEMKLITLSSCSFIVDEHGRVGVVTRDIIISVPSPLCMDSRIDGATFFSAAL

EQAGCVRNGOMFON"

4248..4390

/gene="yz1"

complement(4669..4741)

/gene="trnf"

complement(4669..4741)

/gene="trnf"

/product="tRNA-Phe"

5402..5406

/note="target site duplication"

/rpt_type=direct

5407..11559

repeat_region

/transposon="Tyl/copia-like retrotransposon Machiavelli"

5407..6236

/note="5' LTR"

6239..6252

/note="reverse of sequence is identical to a sequence

within the LTR of Machiavelli"

6369..10163

/gene="gag/pol"

6369..10163

/gene="gag/pol"

/codon_start=1

/product="gag and pol"

/protein_id="AAM22635.1"

/db_xref="GI:20502803"

/translation="MNVFVSKGRPEPLTPEQKEYDHANTMTGAVLSALVRLVYD

ANMOYTDGKOLMDALTXYGSDAGSDILIMESFDIKMVDNRSIYEQHEIQCIAKE

LDHLKVLVDFRVAGCIIAKLPSTWRNFATALKHROEISVENLTASLDVEERAKRD

TSGKGGGSHSANNVQKNHNGKPKSNKPNKTNFKKKNAELTCTFACGEAGHFA

KDCPDADRKGKGNVTASNEEDKGNLPFTFVSQPSMWLDTGANVHVCSDI

NLFSYQGRSDSDVNGSHSVHGTVDLKFTSGKIVQLKNVHHVPSITHKVLVSG

TLLCRDFKFKVLSKLVSKSGOITGKYGCGGLFRESLLDENKNKSNHNCANVDL

ASVHSRLCHINFGSMPLMSLIPNITIVYKSKCHSCVQSKQPKPKHKAERHLA

PLELHSDLCENGVLTGKGRFTMILIDDSRCFYILLKTKDEALDYFKLYKAEVE

VQMLKRLRSRDRGGEFFRVDFDCAEHGIIHERPTTSPESNGIAERKNRTLIDL

NANMLDTCGLSKAMWGEALTSCHVLNRLPMGKEKTPYKWKGRPSLSYLRWGM

AKVNPMTNKRKLGRPTVDFVFGASCSIAVRLVVKSEVDDVYDTIMESRDATFF

EHIPPMKDTHSNRSSEIPIENTPIESFEPQHEIYLEDDNDAPKSKQRVEKSF

GDDFIVLVDDPTTIAFAFASDADDAKWEAHNMDSILNSGTWEITIDRYGCKPVG

CLIVHQMDVKTAFNLGLEDIYMEQDGFVYKQGESKVKLLKLYGLKQAKRQHE

KFTDTLTSAFNAEDRCVYRCGGGVILCLYVDDILIFGTNIDVINKVSLKAS

SFMWKLGEADVILNLIKADGDTLSOSHVVEKVLKRFGECKPSPPTPYPSVTL

RKNKRGIDQLRAYSQIVGSLWLAGATREDISFAVSKLSRPSNPGTDHHAERVMR

YLOKTMISYGIHSGHAVLEGYSDSNWSIDDADELYATSGYFTVGGGAVSWKSKQTI

LTRSTMEAEALADTAIVEAEWRLLEMLDLPVVEKPIPAILMNCQNTVIAKVTSSKD

NGKSSRHVKKRLKSVKRLNSRGSIVSVTYISTDKNLADPFTKGLPRNVEIASRMGR

PE"

```

misc_feature 10706..10714
LTR /note="polypurine tract"
10717..11559
/repeat_region /note="3' LTR"
11560..11564
/repeat_region /note="target site duplication"
13761..13768
/rpt_type=direct
/note="target site duplication"
13769..14246
/rpt_type=direct
complement(13769..14246)
/note="incomplete"
/transposon="retrotransposon-like Ozymandias"
complement(13769..14246)
/note="3' LTR"
complement(14253..14290)
complement(14291..14767)
/note="5' LTR"
14768..14775
/repeat_region /note="target site duplication"
14776..14775
/rpt_type=direct
join(14768..14775,15041,15630..15783)
/note="part of a 1.1-kb tandem duplication interrupted by
the Gnat1 insertion"
15034..15041
/repeat_region /note="target site duplication"
15042..15629
/rpt_type=direct
/transposon="MITE Gnat1"
15042..15053
/repeat_region /note="terminal inverted repeat of Gnat1"
15618..15629
/rpt_type=inverted
/note="terminal inverted repeat of Gnat1"
15630..15637
/rpt_type=direct
/note="target site duplication"
15638..15637
/rpt_type=direct
/note="target site duplication"
BASE COUNT 4482 a 3258 c 3686 g 4357 t
ORIGIN
1 TGTCGACGAGACACAGA 16
|||||
Db 4667 TGTCGACGAGACACAGA 4682

Query Match 64.0%; Score 16; DB 8; Length 15783;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTCGACGAGACACAGA 16
|||||
Db 4667 TGTCGACGAGACACAGA 4682

RESULT 45
HSDJ69001
LOCUS HSDJ69001 21941 bp DNA linear PRI 19-MAR-2001
DEFINITION Human DNA sequence from clone RP4-69001 on chromosome 20 Contains
STSS and GSSs, complete sequence.
ACCESSION AL118521
VERSION AL118521.18 GI:11182354
KEYWORDS HTG.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 21941)
AUTHORS Sehara,H.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
COMMENT On Nov 16, 2000 this sequence version replaced gi:9795208.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission

```

corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; SW, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/Projects/C-elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr20>

IMPORTANT: This sequence is not the entire insert of clone RP4-69001. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone RP5-1123D4 is at 21842 in this sequence. The true right end of clone RP3-409010 is at 100 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP4-69001 is from the library RPCI-4 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2.

```

FEATURES             Location/Qualifiers
     source            1..21941
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /chromosome="20"
                        /clone="RP4-69001"
                        /clone_lib="RPCI-4"
     repeat_region     1..3236
                        /note="LIMA2 repeat: matches 3039..6308 of consensus"
                        3250..3659
                        /note="LIMA9 repeat: matches 5894..6306 of consensus"
                        complement(3573..3797)
                        /note="match: STS: Em:HSPF17G3"
     misc_feature       4058..4253
     repeat_region     /note="MIR repeat: matches 40..248 of consensus"
                        5838..5939
                        /note="L2 repeat: matches 2433..2528 of consensus"
     repeat_region     6071..6302
                        /note="L2 repeat: matches 1965..2200 of consensus"
     repeat_region     6351..6814
                        /note="MT1D repeat: matches 1..505 of consensus"
                        complement(7274..7695)
     misc_feature       /note="match: GSS: Em:AQ432967"
                        complement(7295..7621)
     repeat_region     7442..7661
                        /note="match: GSS: Em:AQ091580"
     repeat_region     8877..9140
                        /note="MIR repeat: matches 21..235 of consensus"
     repeat_region     /note="L2 repeat: matches 2452..2750 of consensus"
                        9385..9444
                        /note="30 copies 2 mer ca 83% conserved"
     repeat_region     9486..9700
                        /note="MER20 repeat: matches 3..217 of consensus"
                        9936..10026
     repeat_region     /note="MIR repeat: matches 52..144 of consensus"
                        complement(10112..10352)
     misc_feature       /note="match: GSS: Em:AQ035997"
                        10833..10958
     repeat_region     /note="L2 repeat: matches 2585..2710 of consensus"
                        11379..11600
                        /note="Charlie2 repeat: matches 3368..3644 of consensus"
                        11601..11726
     repeat_region     /note="L1P4 repeat: matches 5693..5818 of consensus"
                        11737..11880
     repeat_region     /note="L1PAL3 repeat: matches 6013..6156 of consensus"
                        11881..12269

```

```

repeat_region /note="Charlie2 repeat: matches 2944. .3387 of consensus"
12481. .12595
repeat_region /note="Charlie2 repeat: matches 1. .134 of consensus"
12850. .13137
misc_feature /note="L2 repeat: matches 2128. .2507 of consensus"
14108. .14655
misc_feature /note="match: GSS: Em:AQ759431"
14300. .14704
repeat_region /note="match: GSS: Em:A2024068"
17107. .17134
repeat_region /note="7 copies 4 mer gaag 92% conserved"
17367. .17484
repeat_region /note="L2 repeat: matches 2633. .2750 of consensus"
17520. .17621
repeat_region /note="MER5A repeat: matches 1. .113 of consensus"
18393. .18701
misc_feature /note="AluSp repeat: matches 1. .310 of consensus"
complement(18959. .19398)
repeat_region /note="match: GSS: Em:AQ792207"
19444. .19571
repeat_region /note="MER53 repeat: matches 1. .120 of consensus"
19466. .19654
repeat_region /note="MER53 repeat: matches 17. .177 of consensus"
20090. .20313
repeat_region /note="L2 repeat: matches 2471. .2706 of consensus"
20332. .20375
repeat_region /note="11 copies 4 mer atgg 97% conserved"
20430. .20896
repeat_region /note="L1ME2 repeat: matches 5684. .6152 of consensus"
20937. .21156
repeat_region /note="L1MC4 repeat: matches 7765. .7976 of consensus"
BASE COUNT 7244 a 4202 c 4204 g 6291 t
ORIGIN

```

```

Query Match 64.0%; Score 16; DB 9; Length 21941;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 9 GAACCGAATTTACA 24
Db 9708 GAACCGAATTTACA 9723

```

Search completed: November 24, 2002, 14:30:46
Job time : 2396.5 secs

; CURRENT FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 86
; SEQ ID NO 11
; LENGTH: 3609
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION:
; NAME/KEY: unsure
; LOCATION: 92
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 869
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 1385
; OTHER INFORMATION: unknown
US-09-705-299-11

Query Match 4.4%; Score 44; DB 4; Length 3609;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 TGGGFGCTCACACCTGTATCCAGCAGCTTTGGAGGCTGAG 274
|||||
Db 1990 TGGGFGCTCACACCTGTATCCAGCAGCTTTGGAGGCTGAG 2033

RESULT 8

US-09-813-817-3
; Sequence 3, Application US/09813817
; Patent No. 6340583
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001178
; CURRENT APPLICATION NUMBER: US/09/813,817
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 59065
; TYPE: DNA
; ORGANISM: Human
US-09-813-817-3

Query Match 4.4%; Score 44; DB 4; Length 59065;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 CGTCTCTACTAAAAATACAAAATTAGCCAGGTGTGTGGGCACA 377
|||||
Db 58882 CGTCTCTACTAAAAATACAAAATTAGCCAGGTGTGTGGGCACA 58925

RESULT 9

US-09-813-817-3/c
; Sequence 3, Application US/09813817
; Patent No. 6340583
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001178
; CURRENT APPLICATION NUMBER: US/09/813,817
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 59065

; TYPE: DNA
; ORGANISM: Human
US-09-813-817-3

Query Match 4.4%; Score 44; DB 4; Length 59065;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 512 ATAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAA 555
|||||
Db 3447 ATAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAA 3404

RESULT 10

US-09-978-197-3
; Sequence 3, Application US/09978197
; Patent No. 6403353
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001178DIV
; CURRENT APPLICATION NUMBER: US/09/978,197
; CURRENT FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 09/813,817
; PRIOR FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 59065
; TYPE: DNA
; ORGANISM: Human
US-09-978-197-3

Query Match 4.4%; Score 44; DB 4; Length 59065;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 CGTCTCTACTAAAAATACAAAATTAGCCAGGTGTGTGGGCACA 377
|||||
Db 58882 CGTCTCTACTAAAAATACAAAATTAGCCAGGTGTGTGGGCACA 58925

RESULT 11

US-09-978-197-3/c
; Sequence 3, Application US/09978197
; Patent No. 6403353
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001178DIV
; CURRENT APPLICATION NUMBER: US/09/978,197
; CURRENT FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 09/813,817
; PRIOR FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 59065
; TYPE: DNA
; ORGANISM: Human
US-09-978-197-3

Query Match 4.4%; Score 44; DB 4; Length 59065;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 512 ATAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAA 555
|||||
Db 3447 ATAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAA 3404

RESULT 12

US-08-781-891-79/c
; Sequence 79, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,891
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6090620tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 240052.419
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 79:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 87350 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-781-891-79

Query Match 4.4%; Score 44; DB 3; Length 87350;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 354 AAATTAGCAGTGTGGTGGCACACGCGCTGTAGTCCCGACTACI 397
|||||
DB 67204 AAATTAGCAGTGTGGTGGCACACGCGCTGTAGTCCCGACTACI 67161

RESULT 13

US-09-791-211-3/c
; Sequence 3, Application US/09791211
; Patent No. 6448080
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION
; FILE REFERENCE: RTS-0205
; CURRENT APPLICATION NUMBER: US/09/791,211
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 90
; SEQ ID NO 3
; LENGTH: 87543
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure

; LOCATION: 7421
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7427
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 11609
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 12605
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 12742
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 29370
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 29422
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 29979
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 29980
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 29981
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 30136
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 30140
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 31205
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 31206
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 31592
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 33095
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 33160
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 34066
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 34072
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 36816
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 39020
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 42164
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 42459
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 46808
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 46823

OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 46826
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 47291
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 52786
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 52787
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 53384
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 54684
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 59215
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 59235
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 59242
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 63290
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 66614
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68660
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68697
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68718
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68733
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68739
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 69785
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 79134
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 79198
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 86336
OTHER INFORMATION: unknown
OTHER INFORMATION:
US-09-791-211-3

Query Match 4.4%; Score 44; DB 4; Length 87543;
Best Local Similarity 100.0%; Pred. No. 3.8e-12;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 354 AAATTAGCAGGTGTGGTCACAGCGCTGTAGTCCAGCTACT 397
|||||
DB 67397 AAATTAGCAGGTGTGGTCACAGCGCTGTAGTCCAGCTACT 67354

RESULT 14
US-09-593-995-10
; Sequence 10, Application US/09593995
; Patent No. 6406888
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Gao, Zeren
; TITLE OF INVENTION: HELICAL CYTOKINE ZALPHA33
; FILE REFERENCE: 99-38
; CURRENT APPLICATION NUMBER: US/09/593,995
; CURRENT FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: 60/139,121
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 20598
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-593-995-10

Query Match 4.3%; Score 43; DB 4; Length 20598;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 ATAGGCTAGGCGGTGGCTCAGCGCTGTATCCAGCACTTT 127
|||||
DB 3990 ATAGGCTAGGCGGTGGCTCAGCGCTGTATCCAGCACTTT 4032

RESULT 15
US-09-810-671-3/c
; Sequence 3, Application US/09810671
; Patent No. 6455291
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO00758
; CURRENT APPLICATION NUMBER: US/09/810,671
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 21234
; TYPE: DNA
; ORGANISM: Human
US-09-810-671-3

Query Match 4.3%; Score 43; DB 4; Length 21234;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GGGTGGGTGGCTCACACCTGTATCCAGCACTTTGGGAGGC 270
|||||
DB 15212 GGGTGGGTGGCTCACACCTGTATCCAGCACTTTGGGAGGC 15170

RESULT 16
US-09-797-906-3
; Sequence 3, Application US/09797906
; Patent No. 6329188
; GENERAL INFORMATION:
; APPLICANT: Zianghe YAN, Karen A. KETCHUM, Valentina DIFRANCESCO, Ellen M. BEASLEY
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001151CIP
; CURRENT APPLICATION NUMBER: US/09/797,906
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 3
LENGTH: 84495
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(84495)
OTHER INFORMATION: n = A,T,C or G
US-09-797-906-3

Query Match 4.3%; Score 43; DB 4; Length 84495;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 335 GTCTCTACTAAATACAAAAATTAGCCAGGTGTGGTGACACA 377
|||||

DB 3963 GTCTCTACTAAATACAAAAATTAGCCAGGTGTGGTGACACA 4005
|||||

RESULT 17
US-08-441-370-1/c
; Sequence 1, Application US/08441370
; Patent No. 5721138
; GENERAL INFORMATION:
; APPLICANT: Lawn, Richard M.
; TITLE OF INVENTION: Apolipoprotein(A) Promoter and
; TITLE OF INVENTION: Regulatory Sequence Constructs and Methods of Use
; Patent No. 5721138
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/441.370
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/991,849
; FILING DATE: 15-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7627-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2542 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
US-08-441-370-1

Query Match 4.2%; Score 42; DB 1; Length 2542;
Best Local Similarity 100.0%; Pred. No. 3.6e-11;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 511 AATAAATAAATAAATAAATAAATAAATAAATAAATAA 552
|||||

DB 73 AATAAATAAATAAATAAATAAATAAATAAATAAATAA 32
|||||

RESULT 18
US-08-471-454-1
; Sequence 1, Application US/08471454
; Patent No. 5731169
; GENERAL INFORMATION:
; APPLICANT: MOGENSEN, Knud E.
; APPLICANT: UZE, Gilles
; APPLICANT: LUTFALLA, Georges
; APPLICANT: GRESSER, Ion
; TITLE OF INVENTION: CDNA FRAGMENT CODING FOR THE GENE FOR
; TITLE OF INVENTION: THE ALPHA INTERFERON RECEPTOR AND PROCESS FOR THE
; TITLE OF INVENTION: PREPARATION OF THE CORRESPONDING PROTEIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471.454
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/900,642
; FILING DATE: 15-JUN-1992
; APPLICATION NUMBER: FR 89/13770
; FILING DATE: 20-OCT-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 960-7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; TELEX: 200797 NIXN UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2784 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-471-454-1

Query Match 4.2%; Score 42; DB 1; Length 2784;
Best Local Similarity 100.0%; Pred. No. 3.6e-11;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 CGGTGGCTCACACTGTATCCACGACCTTTGGAGGCTGAG 274
|||||

DB 1947 CGGTGGCTCACACTGTATCCACGACCTTTGGAGGCTGAG 1988
|||||

RESULT 19
US-08-466-974-1
; Sequence 1, Application US/08466974
; Patent No. 5861258
; GENERAL INFORMATION:
; APPLICANT: MOGENSEN, Knud E.
; APPLICANT: UZE, Gilles
; APPLICANT: LUTFALLA, Georges
; APPLICANT: GRESSER, Ion
; TITLE OF INVENTION: CDNA FRAGMENT CODING FOR THE GENE FOR
; TITLE OF INVENTION: THE ALPHA INTERFERON RECEPTOR AND PROCESS FOR THE
; TITLE OF INVENTION: PREPARATION OF THE CORRESPONDING PROTEIN

```

? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/471,453
? FILING DATE: 06-JUN-1995
? CLASSIFICATION: 536
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/900,642
? FILING DATE: 15-JUN-1992
? APPLICATION NUMBER: FR 89/13770
? FILING DATE: 20-OCT-1989
? ATTORNEY/AGENT INFORMATION:
? NAME: BYRNE, THOMAS E.
? REGISTRATION NUMBER: 32,205
? REFERENCE/DOCKET NUMBER: 960-7
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (703) 816-4000
? TELEFAX: (703) 816-4100
? TELEX: 200797 NIXN UR
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2784 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? US-08-471-453-1

? Query Match 4.2%; Score 42; DB 2; Length 2784;
? Best Local Similarity 100.0%; Pred No. 3.6e-11;
? Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 233 CGGTGGCTCACACCTGTAATCCACGACTTTGGAGGCTGAG 274
      ||||||||||||||||||||||||||||||||||||||||
Db 1947 CGGTGGCTCACACCTGTAATCCACGACTTTGGAGGCTGAG 1988

RESULT 21
US-09-026-033-17/c
? Sequence 17, Application US/09026033
? Patent No. 6368791
? GENERAL INFORMATION:
? APPLICANT: Felix, Carolyn A.
? APPLICANT: Jones, Douglas H.
? APPLICANT: Rappaport, Eric
? TITLE OF INVENTION: METHOD AND KITS FOR ANALYSIS OF
? TITLE OF INVENTION: CHROMOSOMAL REARRANGEMENTS ASSOCIATED WITH LEUKEMIA
? NUMBER OF SEQUENCES: 23
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
? STREET: One Commerce Square, 2005 Market Street, 22nd
? STREET: Floor
? CITY: Philadelphia
? STATE: PA
? COUNTRY: U.S.A.
? ZIP: 19103-7086
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/026,033
? FILING DATE: 19-FEB-1998
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/038,624
? FILING DATE: 19-FEB-1997
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/056,923
? FILING DATE: 26-AUG-1997
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/065,911
? FILING DATE: 17-NOV-1997
? ATTORNEY/AGENT INFORMATION:

```


NAME: Doyle Leary, Ph.D., Kathryn
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 7600-1001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-965-1284
TELEFAX: 215-567-2991
TELEX: 831-494
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 3441 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-026-033-17

Query Match 4.2%; Score 42; DB 4; Length 3441;
Best Local Similarity 100.0%; Pred. No. 3.6e-11;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 234 GTGGCTCACACCTGTAATCCAGCACCTTTGGGAGGCTGAGA 275
|||||
Db 2330 GTGGCTCACACCTGTAATCCAGCACCTTTGGGAGGCTGAGA 2289

RESULT 22

US-09-345-882-1/c
Sequence 1. Application US/09345882
Patent No. 6399373
GENERAL INFORMATION:
APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
FILE REFERENCE: GENSET-031A
CURRENT APPLICATION NUMBER: US/09/345,882
CURRENT FILING DATE: 1999-06-30
PRIOR APPLICATION NUMBER: US 60/091,315
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/111,909
PRIOR FILING DATE: 1998-12-10
NUMBER OF SEQ ID NOS: 140
SOFTWARE: Patent.Pm
SEQ ID NO 1
LENGTH: 162450
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 72794
OTHER INFORMATION: 5-124-273 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 88073
OTHER INFORMATION: 5-127-261 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 90842
OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 93714
OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
FEATURE:
NAME/KEY: allele
LOCATION: 97122
OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 97152
OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
FEATURE:
NAME/KEY: allele
LOCATION: 99098

OTHER INFORMATION: 5-130-257 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 99117
OTHER INFORMATION: 5-130-276 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 103806
OTHER INFORMATION: 5-131-395 : polymorphic base A or T
FEATURE:
NAME/KEY: allele
LOCATION: 106940
OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108106
OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108149
OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
FEATURE:
NAME/KEY: allele
LOCATION: 108308
OTHER INFORMATION: 5-135-357 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 108471
OTHER INFORMATION: 5-136-174 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134134
OTHER INFORMATION: 5-140-120 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134362
OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 134374
OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
FEATURE:
NAME/KEY: allele
LOCATION: 146328
OTHER INFORMATION: 5-143-84 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 146345
OTHER INFORMATION: 5-143-101 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 150329
OTHER INFORMATION: 5-145-24 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 160031
OTHER INFORMATION: 5-148-352 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52

```
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
FEATURE:
```

```
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
FEATURE:
Query Match 4.2%; Score 42; DB 4; Length 162450;
Best Local Similarity 100.0%; Pred. No. 3.6e-11;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 233 CGTGGCTCACACCTGTATCCCGACACTTTGGGAGGCTGAG 274
Db 78684 CGTGGCTCACACCTGTATCCCGACACTTTGGGAGGCTGAG 78643
RESULT 23
US-09-511-625B-43
; Sequence 43, Application US/09511625B
; Patent No. 6368828
; GENERAL INFORMATION:
; APPLICANT: LaRoche, William J.
; APPLICANT: Patel, Bharvin
; APPLICANT: Pierce, Jacalyn H.
; TITLE OF INVENTION: ATTENUATED AND DOMINANT NEGATIVE VARIANT
; TITLE OF INVENTION: CDNA OF STAT6: STAT6B AND STAT6C
; FILE REFERENCE: 14014.0300u1
; CURRENT APPLICATION NUMBER: US/09/511,625B
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: PCT/US98/17821
; PRIOR FILING DATE: 1998-08-27
; PRIOR APPLICATION NUMBER: 60/070,397
; PRIOR FILING DATE: 1998-01-05
; PRIOR APPLICATION NUMBER: 60/056,075
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 1110
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of artificial sequence:/No. 6368828e -
NAME/KEY: misc feature
LOCATION: (1)..(1110)
; OTHER INFORMATION: No. 6368828e: n = a, t, c, or g
US-09-511-625B-43
Query Match 4.1%; Score 41; DB 4; Length 1110;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 234 GGTGGCTCACACCTGTATCCCGACACTTTGGGAGGCTGAG 274
Db 411 GGTGGCTCACACCTGTATCCCGACACTTTGGGAGGCTGAG 451
RESULT 24
5187077-16/c
; Patent No. 5187077
; APPLICANT: GEARING, DAVID P.; GOUGH, NICHOLAS M.; HILTON,
; DOUGLAS J.; KING, JULIE A.; METCALF, DONALD; NICE, EDOUARD C.
; NICOLA, NICOS A.; SIMPSON, RICHARD J.; WILLSON, TRACY A.
; TITLE OF INVENTION: LEUKEMIA INHIBITORY FACTOR
; NUMBER OF SEQUENCES: 41
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/294,514
; FILING DATE: 09-DEC-1988
; SEQ ID NO:16:
; LENGTH: 1437
```

5187077-16

Query Match 4.1%; Score 41; DB 6; Length 1437;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 512 ATAAATAAAATAAAATAAAATAAAATAAAATAAAATAAA 552
Db 1181 ATAAATAAAATAAAATAAAATAAAATAAAATAAAATAAA 1141

RESULT 25

5427925-14/c
; Patent No. 5427925
; APPLICANT: GEARING, DAVID P.; GOUGH, NICHOLAS M.; HILTON,
; DOUGLAS J.; KING, JULIE A.; METCALF, DONALD; NICE, EDWARD C.
; NICOLA, NICOLO A.; SIMPSON, RICHARD J.; WILSON, TRACY A.
; TITLE OF INVENTION: RECOMBINANT METHOD FOR MAKING
; LEUKEMIA INHIBITOR FACTOR
; NUMBER OF SEQUENCES: 38
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/58,979
; FILING DATE: 06-MAY-1993
; APPLICATION NUMBER: 948,614
; FILING DATE: 22-SEP-1992
; APPLICATION NUMBER: 667,159
; FILING DATE: 11-MAR-1991
; SEQ ID NO: 14:
; LENGTH: 1437
5427925-14

Query Match 4.1%; Score 41; DB 6; Length 1437;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 512 ATAAATAAAATAAAATAAAATAAAATAAAATAAAATAAA 552
Db 1181 ATAAATAAAATAAAATAAAATAAAATAAAATAAAATAAA 1141

RESULT 26

US-08-370-975B-10/c
; Sequence 10, Application US/08370975B
; Patent No. 5622851
; GENERAL INFORMATION:
; APPLICANT: Maley, Frank
; APPLICANT: Maley, Gladys F.
; APPLICANT: Weiner, Karen X.B.
; TITLE OF INVENTION: Human Deoxycytidylate Deaminase Gene
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,975B
; FILING DATE: 10-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Timian, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 20894/80
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716)263-1636
; TELEFAX: (716)263-1600

; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1554 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 4q35
US-08-370-975B-10

Query Match 4.1%; Score 41; DB 1; Length 1554;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 GGTGGCTCACACCTGTAAATCCACGACTTTGGAGGCTGAG 274
Db 636 GGTGGCTCACACCTGTAAATCCACGACTTTGGAGGCTGAG 596

RESULT 27

US-09-078-294-9/c
; Sequence 9, Application US/09078294
; Patent No. 6265211
; GENERAL INFORMATION:
; APPLICANT: Choo, Kong-Hong Andy
; APPLICANT: Du Sart, Desiree
; APPLICANT: Cancilla, Michael R.
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE
; FILE REFERENCE: Davies Col
; CURRENT APPLICATION NUMBER: US/09/078,294
; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1701
; TYPE: DNA
; ORGANISM: BAC-F2 contig 5
US-09-078-294-9

Query Match 4.1%; Score 41; DB 4; Length 1701;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 GGTGGCTCACACCTGTAAATCCACGACTTTGGAGGCTGAG 274
Db 1390 GGTGGCTCACACCTGTAAATCCACGACTTTGGAGGCTGAG 1350

RESULT 28

US-09-880-427-2/c
; Sequence 2, Application US/09880427
; Patent No. 6358728
; GENERAL INFORMATION:
; APPLICANT: Simon, Andras
; APPLICANT: Eriksson, Ulf
; APPLICANT: Dryja, Thaddeus P.
; APPLICANT: Berson, Eliot
; APPLICANT: Yamamoto, Hiroyuji
; TITLE OF INVENTION: Mutations in Nucleic Acid Molecules Encoding 11-Cis Retinol
; TITLE OF INVENTION: Dehydrogenase, The Mutated Proteins, and Uses Thereof
; FILE REFERENCE: IUD 5601
; CURRENT APPLICATION NUMBER: US/09/880,427
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 09/306,538
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 5
; SEQ ID NO 2
; LENGTH: 6330
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure

; LOCATION: 5357, 5448
; OTHER INFORMATION: nucleotide not determined
US-09-880-427-2

Query Match 4.1%; Score 41; DB 4; Length 6330;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 GTGGCTCACACCTGTAATCCAGCAGCTTTGGAGGCTGAG 274

Db 3624 GTGGCTCACACCTGTAATCCAGCAGCTTTGGAGGCTGAG 3584

RESULT 29

US-09-306-538B-2/C

; Sequence 2, Application US/09306538B

; Patent No. 6372463

; GENERAL INFORMATION:

; APPLICANT: Simon, Andras

; APPLICANT: Eriksson, Ulf

; APPLICANT: Dryja, Thaddeus P.

; APPLICANT: Berson, Eliot

; APPLICANT: Yanamoto, Hiroyuji

; TITLE OF INVENTION: Mutations in Nucleic Acid Molecules Encoding 11-Cis Retinol

; TITLE OF INVENTION: Dehydrogenase, The Mutated Proteins, and Uses Thereof

; FILE REFERENCE: LUD 5601

; CURRENT APPLICATION NUMBER: US/09/306,538B

; CURRENT FILING DATE: 1999-05-06

; NUMBER OF SEQ ID NOS: 5

; SEQ ID NO 2

; LENGTH: 6330

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: unsure

; LOCATION: 5357, 5448

; OTHER INFORMATION: nucleotide not determined

US-09-306-538B-2

Query Match 4.1%; Score 41; DB 4; Length 6330;

Best Local Similarity 100.0%; Pred. No. 1.1e-10;

Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 GTGGCTCACACCTGTAATCCAGCAGCTTTGGAGGCTGAG 274

Db 3624 GTGGCTCACACCTGTAATCCAGCAGCTTTGGAGGCTGAG 3584

RESULT 30

US-08-646-301A-1

; Sequence 1, Application US/08646301A

; Patent No. 6194211

; GENERAL INFORMATION:

; APPLICANT: Richards, Cynthia Ann

; APPLICANT: Huber, Brian E.

; TITLE OF INVENTION: Transcriptional Regulatory Sequence of Carcinoembryonic

; Patent No. 6194211

; TITLE OF INVENTION: Antigen for Expression Targeting

; FILE REFERENCE: PB1508USW

; CURRENT APPLICATION NUMBER: US/08/646,301A

; CURRENT FILING DATE: 1996-05-16

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 11288

; TYPE: DNA

; ORGANISM: Homo sapiens

US-08-646-301A-1

Query Match

Best Local Similarity 100.0%; Pred. No. 1.1e-10;

Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 GCGCGGTGGCTCAGCCTGTAATCCAGCAGCTTTAGAAGG 133
|||||
Db 540 GCGCGGTGGCTCAGCCTGTAATCCAGCAGCTTTAGAAGG 580

RESULT 31

US-08-481-968A-4

; Sequence 4, Application US/08481968A

; Patent No. 6300490

; GENERAL INFORMATION:

; APPLICANT: Huber, Brian

; APPLICANT: Richards, Cynthia

; TITLE OF INVENTION: Molecular Constructs Comprising a Carcinoembryonic Antigen (

; TITLE OF INVENTION: Transcriptional Regulatory Region

; FILE REFERENCE: PB1087US4

; CURRENT APPLICATION NUMBER: US/08/481,968A

; CURRENT FILING DATE: 1998-06-07

; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 4

; LENGTH: 11288

; TYPE: DNA

; ORGANISM: Homo sapiens

US-08-481-968A-4

Query Match 4.1%; Score 41; DB 4; Length 11288;

Best Local Similarity 100.0%; Pred. No. 1.1e-10;

Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 GCGCGGTGGCTCAGCCTGTAATCCAGCAGCTTTAGAAGG 133

Db 540 GCGCGGTGGCTCAGCCTGTAATCCAGCAGCTTTAGAAGG 580

RESULT 32

US-08-154-712B-4

; Sequence 4, Application US/08154712B

; Patent No. 6337209

; GENERAL INFORMATION:

; APPLICANT: Huber, Brian

; APPLICANT: Richards, Cynthia

; TITLE OF INVENTION: Molecular Constructs Containing a Carcinoembryonic Antigen R

; TITLE OF INVENTION: Sequence

; FILE REFERENCE: PB1087US3

; CURRENT APPLICATION NUMBER: US/08/154,712B

; CURRENT FILING DATE: 1993-11-19

; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 4

; LENGTH: 11288

; TYPE: DNA

; ORGANISM: Homo sapiens

US-08-154-712B-4

Query Match 4.1%; Score 41; DB 4; Length 11288;

Best Local Similarity 100.0%; Pred. No. 1.1e-10;

Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 GCGCGGTGGCTCAGCCTGTAATCCAGCAGCTTTAGAAGG 133

Db 540 GCGCGGTGGCTCAGCCTGTAATCCAGCAGCTTTAGAAGG 580

RESULT 33

US-09-801-052-3

; Sequence 3, Application US/09801052

; Patent No. 6368842

; GENERAL INFORMATION:

; APPLICANT: BEASLEY, Ellen

; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHOLIPASE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHOLIPASE

; TITLE OF INVENTION: PROTEINS, AND USES THEREOF

; FILE REFERENCE: CL001045

; CURRENT APPLICATION NUMBER: US/09/801,052
; CURRENT FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 16063
; TYPE: DNA
; ORGANISM: Human
US-09-801-052-3

Query Match 4.1%; Score 41; DB 4; Length 16063;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 GTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAG 274
|||||
Db 10954 GTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAG 10994

RESULT 34
US-09-511-625B-5
; Sequence 5, Application US/09511625B
; Patent No. 6368828
; GENERAL INFORMATION:
; APPLICANT: LaRocheville, William J.
; APPLICANT: Patel, Bhavvin
; APPLICANT: Pierce, Jagalyn H.
; TITLE OF INVENTION: ATTENUATED AND DOMINANT NEGATIVE VARIANT
; TITLE OF INVENTION: CDNAS OF STAT6; STAT6B AND STAT6C
; FILE REFERENCE: 14014.0300ul
; CURRENT APPLICATION NUMBER: US/09/511.625B
; CURRENT FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: PCT/US98/17821
; PRIOR FILING DATE: 1998-08-27
; PRIOR APPLICATION NUMBER: 60/070,397
; PRIOR FILING DATE: 1998-01-05
; PRIOR APPLICATION NUMBER: 60/056,075
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 17425
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of artificial Sequence./No. 6368828e -
; NAME/KEY: misc.feature
; LOCATION: (1)...(17425)
; OTHER INFORMATION: n = a, t, c or g
US-09-511-625B-5

Query Match 4.1%; Score 41; DB 4; Length 17425;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 GTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAG 274
|||||
Db 11733 GTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAG 11773

RESULT 35
US-08-370-975B-1/c
; Sequence 1, Application US/08370975B
; Patent No. 5622851
; GENERAL INFORMATION:
; APPLICANT: Maley, Frank
; APPLICANT: Maley, Gladys F.
; APPLICANT: Weiner, Karen X.B.
; TITLE OF INVENTION: Human Deoxycytidylate Deaminase Gene
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle

; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,975B
; FILING DATE: 10-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Timian, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 20894/80
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716)263-1636
; TELEFAX: (716)263-1600
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26764 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 4q35
US-08-370-975B-1

Query Match 4.1%; Score 41; DB 1; Length 26764;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 GTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAG 274
|||||
Db 24473 GTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAG 24433

RESULT 36
US-08-965-048-5/c
; Sequence 5, Application US/08965048
; Patent No. 6323244
; GENERAL INFORMATION:
; APPLICANT: Chen, Hong
; APPLICANT: Freimer, Nelson
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF NEUROPSYCHIATRIC DISORDERS
; FILE REFERENCE: 7853-093
; CURRENT APPLICATION NUMBER: US/08/965,048
; CURRENT FILING DATE: 1997-11-05
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 45716
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-965-048-5

Query Match 4.1%; Score 41; DB 4; Length 45716;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 GTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAG 274
|||||
Db 20736 GTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAG 20696

RESULT 37
US-08-965-048-6/c
; Sequence 6, Application US/08965048

```
; Patent No. 6323244
; GENERAL INFORMATION:
; APPLICANT: Chen, Hong
; APPLICANT: Freimer, Nelson
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF NEUROPSYCHIATRIC DISORDERS
; FILE REFERENCE: 7853-093
; CURRENT APPLICATION NUMBER: US/08/965,048
; CURRENT FILING DATE: 1997-11-05
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 45989
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: n=a, c, g, or t
US-08-965-048-6

Query Match      4.1%; Score 41; DB 4; Length 45989;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 GTGGCTCACACCTGTATCCAGCAGCTTTGGAGGCTGAG 274
Db 20850 GTGGCTCACACCTGTATCCAGCAGCTTTGGAGGCTGAG 20810

RESULT 38
US-09-784-316-3
; Sequence 3, Application US/09784316
; Patent No. 6461843
; GENERAL INFORMATION:
; APPLICANT: Wei, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001139
; CURRENT APPLICATION NUMBER: US/09/784,316
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 65042
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(65042)
; OTHER INFORMATION: n = A,T,C or G
US-09-784-316-3

Query Match      4.1%; Score 41; DB 4; Length 65042;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 GTGGCTCACACCTGTATCCAGCAGCTTTGGAGGCTGAG 274
Db 13735 GTGGCTCACACCTGTATCCAGCAGCTTTGGAGGCTGAG 13775

RESULT 39
US-09-268-992-7/c
; Sequence 7, Application US/09268992
; Patent No. 6342351
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Freimer, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/268,992
; CURRENT FILING DATE: 1999-03-16
; EARLIER FILING DATE: 1999-03-16
; EARLIER FILING DATE: 1999-01-22
```

```
; EARLIER APPLICATION NUMBER: 60/106,056
; EARLIER FILING DATE: 1998-10-28
; EARLIER APPLICATION NUMBER: 60/088,312
; EARLIER FILING DATE: 1998-06-05
; EARLIER APPLICATION NUMBER: 60/078,044
; EARLIER FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 72604
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g, or t
US-09-268-992-7

Query Match      4.1%; Score 41; DB 4; Length 72604;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 GTGCGTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGC 270
Db 2857 GTGCGTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGC 2817

RESULT 40
US-09-657-474-7/c
; Sequence 7, Application US/09657474
; Patent No. 6399762
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Freimer, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/657,474
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/268,992
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/236,134
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/106,056
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/088,312
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/078,044
; PRIOR FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 72604
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g, or t
US-09-657-474-7

Query Match      4.1%; Score 41; DB 4; Length 72604;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 GTGCGTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGC 270
Db 2857 GTGCGTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGC 2817

RESULT 41
US-09-741-150-3
; Sequence 3, Application US/09741150
```

```
; Patent No. 6436689
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000968
; CURRENT APPLICATION NUMBER: US/09/741.150
; CURRENT FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 112132
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(112132)
; OTHER INFORMATION: n = A,T,C or G
US-09-741-150-3

Query Match          4.1%; Score 41; DB 4; Length 112132;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 GGTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAG 274
|||||
Db 105191 GGTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAG 105231

RESULT 42
US-08-724-394A-20
; Sequence 20, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Krommal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; TITLE OF INVENTION: Sequences and Antibodies Thereto
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-21

Query Match          4.1%; Score 41; DB 2; Length 246240;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-20

Query Match          4.1%; Score 41; DB 2; Length 246240;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 GGTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAG 274
|||||
Db 7967 GGTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAG 8007

RESULT 43
US-08-724-394A-21
; Sequence 21, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Krommal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; TITLE OF INVENTION: Sequences and Antibodies Thereto
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-21

Query Match          4.1%; Score 41; DB 2; Length 246240;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 GGTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAG 274
```

```
Db 7967 GTGGCTCACACCTGTAATCCAGCAGCTTTGGGAGGCTGAG 8007
|||||
; TITLE OF INVENTION: AORTIC STENOSIS AND WILLIAMS SYNDROME
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
; STREET: 1201 New York Ave., N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,020A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 08/041,576
; FILING DATE: 05-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 19780-105509
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4800
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1300 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-08-474-020A-13

Query Match 4.0%; Score 40; DB 2; Length 1300;
Best Local Similarity 100.0%; Pred. No. 3.3e-10;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GTGGCTCACACCTGTAATCCAGCAGCTTTGGGAGGCTGAG 274
|||||
Db 414 GTGGCTCACACCTGTAATCCAGCAGCTTTGGGAGGCTGAG 375

Search completed: November 24, 2002, 12:19:13
Job time : 1108 secs
```

```
Db 7967 GTGGCTCACACCTGTAATCCAGCAGCTTTGGGAGGCTGAG 8007
|||||
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
; US-08-724-394A-22

Query Match 4.1%; Score 41; DB 2; Length 246240;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 GTGGCTCACACCTGTAATCCAGCAGCTTTGGGAGGCTGAG 274
|||||
Db 7967 GTGGCTCACACCTGTAATCCAGCAGCTTTGGGAGGCTGAG 8007

RESULT 45
US-08-474-020A-13/c
; Sequence 13, Application US/08474020A
; Patent No. 5840489
; GENERAL INFORMATION:
; APPLICANT: Keating, Mark T.
; APPLICANT: Leppert, Mark F.
; APPLICANT: Morris, Colleen A.
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF SUPRAVALVULAR
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2002, 08:57:23 ; Search time 101 Seconds
(without alignments)
3749.805 Million cell updates/sec

Title: US-09-784-423-32
Perfect score: 1000
Sequence: 1 GGTGTGACCTTATCTCTCT.....CAGAATGATGACGGCGTGC 1000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues
Total number of hits satisfying chosen parameters: 671156

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	999	99.9	1000	10	US-09-784-423-32	Sequence 32, Appl
2	320.2	32.0	7017	10	US-09-764-877-3773	Sequence 3773, Ap
3	320.2	32.0	20522	10	US-09-764-877-3774	Sequence 3774, Ap
4	307	30.7	17424	10	US-09-967-768A-314	Sequence 314, App
5	301.4	30.1	17424	10	US-09-967-768A-314	Sequence 314, App
6	300.6	30.1	32816	10	US-09-729-094-3	Sequence 3, Appli
7	297.2	29.7	24533	9	US-09-764-868-1349	Sequence 1349, Ap
8	296	29.6	15297	9	US-10-003-295-3	Sequence 3, Appli
9	293.8	29.4	30350	9	US-10-118-328-3	Sequence 3, Appli
10	293.8	29.4	99014	10	US-09-880-107-3428	Sequence 3428, Ap
11	291.6	29.2	52216	10	US-09-747-810-1	Sequence 1, Appli
12	291.2	29.1	40645	10	US-09-818-656A-3	Sequence 3, Appli
13	290.8	29.1	46527	10	US-09-933-267A-1	Sequence 1, Appli
14	289.8	29.0	3273	10	US-09-764-878-363	Sequence 363, App
15	289.8	29.0	3276	10	US-09-764-878-362	Sequence 362, App
16	285	28.5	11617	9	US-09-860-670-265	Sequence 265, App
17	284	28.4	2713	10	US-09-967-736-6	Sequence 6, Appli
18	284	28.4	7720	10	US-09-954-456-946	Sequence 946, App
19	284	28.4	7720	10	US-09-954-456-1589	Sequence 1589, Ap

20	282.2	28.2	16086	10	US-09-764-877-2385	Sequence 2385, Ap
21	281.8	28.2	14176	10	US-09-764-864-1644	Sequence 1644, Ap
22	281	28.1	15016	10	US-09-890-107-3783	Sequence 3783, Ap
23	280.4	28.0	33239	10	US-09-814-950-3	Sequence 3, Appli
24	280.2	28.0	32132	10	US-09-764-877-2308	Sequence 2308, Ap
25	279.6	28.0	53332	10	US-09-801-861-3	Sequence 3, Appli
26	279.2	27.9	31168	9	US-09-764-868-1464	Sequence 1464, Ap
27	279.2	27.9	32193	10	US-09-764-877-2623	Sequence 2623, Ap
28	277.8	27.8	22484	10	US-09-875-114-2	Sequence 2, Appli
29	277.8	27.8	22484	10	US-09-880-107-3341	Sequence 3341, Ap
30	276.6	27.7	16086	10	US-09-764-877-2385	Sequence 2385, Ap
31	276.4	27.6	56737	10	US-09-782-378A-17	Sequence 17, Appl
32	276.4	27.6	84539	10	US-09-962-436-36	Sequence 36, Appl
33	276.2	27.6	145831	10	US-09-969-708-79	Sequence 79, Appl
34	276.2	27.6	145831	10	US-09-954-456-2116	Sequence 2116, Ap
35	275.2	27.5	8701	10	US-09-764-847-1976	Sequence 1976, Ap
36	275.2	27.5	11821	10	US-09-764-877-2857	Sequence 2857, Ap
37	274.6	27.5	63000	10	US-09-780-172-18	Sequence 18, Appl
38	274.4	27.4	32193	10	US-09-764-877-2147	Sequence 2147, Ap
39	274	27.4	4963	10	US-09-764-877-2903	Sequence 2903, Ap
40	273.8	27.4	27359	10	US-09-822-863-3	Sequence 3, Appli
41	272.6	27.3	29629	12	US-10-135-689-3	Sequence 3, Appli
42	272.4	27.2	2892	10	US-09-764-878-197	Sequence 197, App
43	272.4	27.2	15275	10	US-09-764-869-1475	Sequence 1475, Ap
44	272.2	27.2	29695	10	US-09-752-820A-3	Sequence 3, Appli
45	272.2	27.2	29695	10	US-09-813-319A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-784-423-32
; Sequence 32, Application US/09784423
; Patent No. US20020012924A1
; GENERAL INFORMATION:
; APPLICANT: Schumm, James W.
; Bacher, Jeffery W.
; TITLE OF INVENTION: MATERIALS AND METHODS FOR IDENTIFYING AND ANALYZING INTERMEDIATE TANDEM REPEAT DNA MARKERS
; NUMBER OF SEQUENCES: 147
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Promega Corporation
; STREET: 2800 Woods Hollow Road
; CITY: Madison
; STATE: Wisconsin
; COUNTRY: U.S.A.
; ZIP: 53711-5399
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Mb
; COMPUTER: IBM compatible PC
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Word 97 (DOS text format)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/784,423
; FILING DATE: 15-Feb-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/018,584
; FILING DATE: 04-Feb-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Grady J. Frenchick
; REGISTRATION NUMBER: 29,018
; REFERENCE/DOCKET NUMBER: 16026.9180
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 257-3501
; TELEFAX: (608) 257-2275
; INFORMATION FOR SEQ ID NO: 32
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1000 bp
; TYPE: Nucleic Acid
; STRANDEDNESS: Double

TOPOLOGY: Circular
MOLECULE TYPE: Genomic DNA
HYPOTHETICAL: no
IMMEDIATE SOURCE:
CLONE: SL32
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 22
SEQUENCE DESCRIPTION: SEQ ID NO: 32

US-09-784-423-32

Query Match 99.9%; Score 999; DB 10; Length 1000;
Best Local Similarity 100.0%; Pred. No. 1.7e-173;
Matches 1000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGTGACCTTATCTCTGACCTCAGTTTCCATCCGTAATAAAGAGCTGTAG 60
DB 1 GGTGTGACCTTATCTCTGACCTCAGTTTCCATCCGTAATAAAGAGCTGTAG 60

QY 61 ATTGTGTAAAAAATAAATGAATAGGCTAGGCGGCTGCTCAGCGCTGTATCCCA 120
DB 61 ATTGTGTAAAAAATAAATGAATAGGCTAGGCGGCTGCTCAGCGCTGTATCCCA 120

QY 121 GCACCTTAGAGTGAAGAGGTGGATCAGTTCAGGTCAGGAGTTTGAGACAGCGTG 180
DB 121 GCACCTTAGAGTGAAGAGGTGGATCAGTTCAGGTCAGGAGTTTGAGACAGCGTG 180

QY 181 GCACACGCGTGAACCCCATCTCTACTAAAAATAAAAAATAGCTNGGTCGGTGGCT 240
DB 181 GCACACGCGTGAACCCCATCTCTACTAAAAATAAAAAATAGCTNGGTCGGTGGCT 240

QY 241 CACACCTGTATCCCGACACTTTGGAGGCTGAGACGGTGGATCAGCTGAAGTCAGGAG 300
DB 241 CACACCTGTATCCCGACACTTTGGAGGCTGAGACGGTGGATCAGCTGAAGTCAGGAG 300

QY 301 TTCAGGCGAGCTGGGCAACATGGTGAACACAGCTCTCTACTAAAAATAAAAAATTAG 360
DB 301 TTCAGGCGAGCTGGGCAACATGGTGAACACAGCTCTCTACTAAAAATAAAAAATTAG 360

QY 361 CCAGTGTGGTGGCAGACAGCTCTAGTCCAGCTACTTGGAGGCTGAGCGGGAAGATC 420
DB 361 CCAGTGTGGTGGCAGACAGCTCTAGTCCAGCTACTTGGAGGCTGAGCGGGAAGATC 420

QY 421 GCTTGAACCCAGTAGGAGGTTGCAGTGAAGGAGTGAAGATCAGTCCAGCTCCAGCC 480
DB 421 GCTTGAACCCAGTAGGAGGTTGCAGTGAAGGAGTGAAGATCAGTCCAGCTCCAGCC 480

QY 481 TGGGTGACAGAGCAAGCTCCCTCTCAGAAATAAATAAATAAATAAATAAATAA 540
DB 481 TGGGTGACAGAGCAAGCTCCCTCTCAGAAATAAATAAATAAATAAATAAATAA 540

QY 541 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 600
DB 541 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 600

QY 601 AGTAATAGCTATCAATATCCACCCCTACCTGCTGCTGAATTTAGTTCTTTTGTG 660
DB 601 AGTAATAGCTATCAATATCCACCCCTACCTGCTGCTGAATTTAGTTCTTTTGTG 660

QY 661 ACCCCCAATAGCTTAAGCAGAAATTCACCGCTACTCTCTGTAATTTCTGTTGCT 720
DB 661 ACCCCCAATAGCTTAAGCAGAAATTCACCGCTACTCTCTGTAATTTCTGTTGCT 720

QY 721 GGCACATAGTTGGTCTCAGTGAAACATGGTGAAGTGAATGAGCAAAATGCAAGAAATCC 780
DB 721 GGCACATAGTTGGTCTCAGTGAAACATGGTGAAGTGAATGAGCAAAATGCAAGAAATCC 780

QY 781 AGCCCATCTGGAGGCCCTCCAGGCGGCTGAGTTCGGGAACTCAATAGTCTGCTCAAT 840
DB 781 AGCCCATCTGGAGGCCCTCCAGGCGGCTGAGTTCGGGAACTCAATAGTCTGCTCAAT 840

QY 841 GGCCCACTGAAGCTAGAGAGTTCTGGGTCCTCCAGCTCCGACCCCATCTCCGACTCAC 900
DB 841 GGCCCACTGAAGAGTTCTGGGTCCTCCAGCTCCGACCCCATCTCCGACTCAC 900

QY 901 TGTGAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 960
DB 901 TGTGAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 960

QY 961 CAGGACTCAGAGGAGCCCGAGAGATGATGACCGGCGTGC 1000
DB 961 CAGGACTCAGAGGAGCCCGAGAGATGATGACCGGCGTGC 1000

RESULT 2

US-09-764-877-3773/c
; Sequence 3773, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 3773
; LENGTH: 7017
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3773

Query Match 32.0%; Score 320.2; DB 10; Length 7017;

Best Local Similarity 84.8%; Pred. No. 4.1e-50;

Matches 406; Conservative 0; Mismatches 69; Indels 5; Gaps 4;

QY 86 TAGGCTAGGCGGGTGGCTCAGCGCTGTAATCCAGCAGCTTTAGAGGTCGAAGAGGTTG 145
DB 4509 TGGGCGACAGCGGGTGGCTCAGCGCTGTAATCCAGCAGCTTTGAGAGCGGAGGTTG 4450

QY 146 GATCACTTTGAGTCAAGGAGTTTGGAGCCAGCGCTGGCCACACGGTGAACCCCATCTCT 205
DB 4449 GATCACTTGAGCAAGAG--TTGAGAGCAGCGCTGGCCACACATGGTGAACCTCGTCTCT 4391

QY 206 ACTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 265
DB 4390 ACTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4331

QY 266 GAGGCTGAGACGGGTGGATCAGCTGAAGTCAAGGAGTTCAAGGCCAGCTGGGCAACATGG 325
DB 4330 GAGGCCAAGCGCGGAGATCA--TGAGGTCAGAGATGGAGACCATCTCTGCTTAACATGG 4273

QY 326 TGAACCCAGCTCTCTACTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 385
DB 4272 TGAACCCAGCTCTCTACTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4213

QY 386 GTCCAGCTACTTTGGAGGCTGAGCGGGAAGATCGCTTGAACCCAGTAGGAGAGTTG 445
DB 4212 GTCCAGCTACTTTGGAGGCTGAGCGGAGATGAGGAGAGATGAGGAGGAGGAGTTG 4153

QY 446 CAGTGAAGCGGAGATGAAGTCACTGCACTCCAGCTGGGTGACAGAGCAAGACTCCCTCT 505
DB 4152 CAGTGAAGCGGAGATCA--CGACACTGCACTCCAGCTGGGCAACAGAGGAGAAATCCGTCT 4094

QY 506 CAGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 565
DB 4093 CA-AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4035

RESULT 3

US-09-764-877-3774/c
; Sequence 3774, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

```
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3774
; LENGTH: 20522
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3774

Query Match      32.0%; Score 320.2; DB 10; Length 20522;
Best Local Similarity 84.6%; Pred. No. 4.5e-50;
Matches 406; Conservative 0; Mismatches 69; Indels 5; Gaps 4;

QY 86 TAGGCTAGGCGCGGTGGCTCAGCGCTGTAATCCAGCAGCTTTAGAGGTCGAAGAGGGTG 145
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10327 TGGGCCAGACGCGGTGGCTCAGCGCTGTAATCCAGCAGCTTTGGAAGCGCGAGGGGTG 10268

QY 146 GATCACTGTAGGTCAGGAGTTTGAGACAGCGCTGGCCAAACACAGGTCAACCCCACTCTCT 205
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10267 GATCACTGTAGGTCAGGAGTTTGAGACAGCGCTGGCCAAACAGGTCAACCCCACTCTCTCT 10209

QY 206 ACTAAATAATAAAATTTAGCTNGGGTGGCTGGCTCAGACCTGTAAATCCAGCAGCTTTGG 265
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10208 ACTAAATAATAAAATTTAGCTNGGGTGGCTGGCTCAGACCTGTAAATCCAGCAGCTTTGG 10149

QY 266 GAGGCTGAGACGGGTGGATCACTGGAAGTCAGGAGTCAAGCCAGCGCTGGGCAACATGG 325
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10148 GAGGCAAGCGCGCGAGATCA-TGAGGTGAGAAGATGGAGACCATCTCTGGCTAACATGG 10091

QY 326 TGAACACAGCTCTCTACTAAATAACAAAATTTAGCCAGGTGTGGTGGCACAGCGCTGTA 385
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10090 TGAACACAGCTCTCTACTAAATAACAAAATTTAGCCAGGTGTGGTGGCACAGCGCTGTA 10031

QY 386 GTCCAGCAGTCTCTGGGAGGCTGAGCGGAAGATCGCTTTGAACCCAGTAGGAGAGGTTG 445
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10030 GTCCAGCAGTCTCTGGGAGGCTGAGCGGAAGATCGCTTTGAACCCAGTAGGAGAGGTTG 9971

QY 446 CAGTGAGCGGAGATAGAGTCACTGCATCCAGCTCCAGCTGGGTGACAGCAAGACTCCCTCT 505
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9970 CAGTGAGCGGAGATCA-CGACACTGCATCCAGCTCCAGCTGGGTGACAGCAAGACTCCCTCT 9912

QY 506 CAGAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 565
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9911 CA-AAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 9853

RESULT 4
US-09-967-768A-314/C
; Sequence 314, Application US/09967768A
; Patent No. US20020150877A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE OF INVENTION: Sets
; FILE REFERENCE: 689290-72
; CURRENT APPLICATION NUMBER: US/09/967,768A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/60/236,109
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,034
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,111
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 314
; LENGTH: 174424
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-967-768A-314
```

```
Query Match      30.7%; Score 307; DB 10; Length 174424;
Best Local Similarity 83.1%; Pred. No. 1.4e-47;
Matches 409; Conservative 0; Mismatches 76; Indels 7; Gaps 5;

QY 76 TTAATGGAATAGGCTAGGCGCGGTGGCTCAGCGCTGTAATCCAGCAGCTTTAGAGGTC 135
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 132168 TAAAAAATAAAGCTAGGCGCAGTGGCTCAGCGCTGTAATCCAGCAGCTTTGGAGGCG 132109

QY 136 GAAGAGGTGATCAGCTTGGAGTCAGGAGTTTGAGACAGCGCTGGCCAAACAGCGTGA 195
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 132108 GAGGCGAGTGGATCATCTGAGTTCAGGAG-TTTCAGACAGCGCTAGCCATGTGGTGAA 132050

QY 196 CCCATCTCTCTACTAAAAAT---AAAAAATTTAGCTNGGGTGGCTGGCTGCATCAGCTGTAAT 252
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 132049 CCCGCTCTCTACCAAAAATACAAAAAATTTAGCCGCGCAGCGTGGCTCAGCGCTAAT 131990

QY 253 CCCAGCAGCTTTGGAGGCTGAGACGCGTGGATCACCTGGAAGTCAGGAGTTTCAGGCGCAGC 312
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 131989 CCCAACACTTTGGGAGCGAGCGCGGTGGATCAGTTGAGGTTCAGGAGTTTCAGACCGAGT 131930

QY 313 CTGGGCAACATGGTGAACCGCTCTCTACTAAAAA-ATACAAAAATTTAGCCAGGTGGT 371
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 131929 CTGGGCAACATGGTGAACCGCTCTCTACTAAAAATTTAGCCAGCGCTGGT 131870

QY 372 GGCACAGCGCTGTAGTCCAGCTTACCTGGAGGCTGAGCGGAAGAATTCGTTGAACCCA 431
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 131869 GGCAGCGCTGTATCCAGCTTACTTGGAGGCTGAGACGGGAGAGATTGCTTGAACCCA 131810

QY 432 GTAGGCGAGGTTGAGTGGCGGAGATGAAGTCACTGCATCCAGCTGGGTGACAGA 491
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 131809 GGAGGCGAGGTTGAGTGGCGGAGAT-CGTGCCATTGCTCCAGCTGGCGGCGAGA 131751

QY 492 GCAGAGCTCCCTCTCAG-AAAATAATAATAATAATAATAATAATAATAATAATAATA 550
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 131750 GTGAGACTTCTCTCAGAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAA 131691

QY 551 AATAAAATTTCTA 562
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 131690 AACAAAAATTA 131679
```

```
RESULT 5
US-09-967-768A-314
; Sequence 314, Application US/09967768A
; Patent No. US20020150877A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE OF INVENTION: Sets
; FILE REFERENCE: 689290-72
; CURRENT APPLICATION NUMBER: US/09/967,768A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/60/236,109
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,034
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,111
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 314
; LENGTH: 174424
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-967-768A-314
```

```
Query Match      30.1%; Score 301.4; DB 10; Length 174424;
Best Local Similarity 81.7%; Pred. No. 1.4e-46;
Matches 397; Conservative 0; Mismatches 82; Indels 7; Gaps 4;

QY 70 AAAAAATTAATGATAGGCTAGGCGCGGTGGCTCAGCGCTGTAATCCAGCAGCTTAG 129
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 69715 AAAACAAAAACATTTTCAGCGCGGCATCGGTGGCTCAGCGCTGTAATCCCAAGCACTTTGG 69774
QY 130 AAGCTCGAAGAGGGTGGATCACTTGAGGTCAGGAGTTTTCAGACCAAGCTCGCCCAACACG 189
Db 69775 GAGGCTGAGGCAAGCAGATCACCTGAAGTCAGGAG-TTCGAGCAAGCCTGGCCACATG 69833
QY 190 GTGAACCCCATCTCTACTATAAATAAAAAATTA---GCTNGGGTGGGCTGCCTCACACC 246
Db 69834 GTGAACCCCTGTCTCTACTATAAATAAAAAATFAAAGGGTCGGGCACGCTGGCCACACC 69893
QY 247 TGTATATCCAGACACTTTGGGAGGCTGAGACGGGTGGATCAGCTGAAGTCAGGAGTTCAAG 306
Db 69894 TGTATATCCAGCACTTTGGGAGCGCCAGGCGGATCAGCTGAGGTCAGGAGTTCAAG 69953
QY 307 GCGAGCTGGGCAACATGTCGAACCAACGCTCTACTATAAATAAAAAATTAAGCCAGGT 366
Db 69954 ACCAGCTTGACAAATATGTTGAACCCCTGTCTCTACTATAAATAAAAAATTCAGCCAGGC 70013
QY 367 GTGGTGGCACACGCCCTGTAGTCCAGCTACTTTGGGAGGCTGAGCGGGAAGATCGCTTGA 426
Db 70014 GTGGTGGCGGCACCTGTAGTCCAGCTACTCAGGAGGCTGAGCAGAGATTCGCTTGA 70073
QY 427 ACCAGTAGCAGAGTTGCAAGTGCAGGAGATAGAGTCACTGCACCTCCAGCCTGGGTG 486
Db 70074 ACCTGGGAGCAGAGTTGCAGTGAGCGGACAT-TGCGCCACTGCACCTAGCCTGGGCG 70132
QY 487 ACAGAGCAAGACTCCCTCTCAGAAAAATAAATAAAAAATAAATAAAAAATAAATAA 546
Db 70133 ACAGAGAAAGTCTCTCAAAAAATAA--TAATAAATAAATAAATAAATAAATAA 70190
QY 547 ATAAAA 552
Db 70191 TAAAAA 70196

RESULT 6
US-09-729-094-3
; Sequence 3, Application US/09729094
; Patent No. US2002019028A1
; GENERAL INFORMATION:
; APPLICANT: CHATURVEDI, Kabir et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; FILE REFERENCE: AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/729,094
; CURRENT FILING DATE: 2000-12-05
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 32816
; TYPE: DNA
; ORGANISM: Human
US-09-729-094-3

Query Match 30.1%; Score 300.6; DB 10; Length 32816;
Best Local Similarity 80.3%; Pred. No. 1.7e-46;
Matches 376; Conservative 0; Mismatches 90; Indels 2; Gaps 2;
QY 85 ATAGCTAGGCGGGTGGCTCAGCGCTGTAATCCAGCACTTTAGAGGTCGAAGAGGT 144
Db 16124 ACAGCGCGGCGGCTGAGGCTCACCTGTATCCCAACATTTTGGAGGCGGAGTGGGC 16183
QY 145 GATCACTTAGGTCAGGAGTTTTCAGACCAAGCCTGGCCACACAGGTCGAACCCCATCTC 204
Db 16184 AGATCGCTAGGTCAGGAG-TTTTGAACCAAGCCTGGCCACATGTTGAACCCGCTCTC 16242
QY 205 TACTAAAAATAAAATTAAGCTNGGGTGGCTGGCTCAACCTGTAAATCCAGCACTTTG 264
Db 16243 TACTAAAAATAAAATTAAGCGGGCGTGGTAGCCCAAGCCTGAAATCCAGCACTTTG 16302
QY 265 GGAGGCTGAGCGGGTGGATCAGCTGAAGTCAGGAGTTCAAGGCCAGCCTGGGCAACATG 324

Db 16303 GGAGGCCAAGAGCGGCGGATCCCTGAGGTCAGAGCTCGAGACCCAGCCTGGCCACATG 16362
QY 325 GTCAAAACCACTCTCTACTATAAATAACAAAAATAGCCAGGTGGTGGCCACACGCTGT 384
Db 16363 GTGAACCCCATCTCTACTATAAATAACAAAAATAGCTGGGCATGCTGGCACATGCTGT 16422
QY 385 AOTCCCACTACTTTGGGAGGCTGAGCGCGGAGAAATCGCTTGAACCCAGTAGCAGAGTT 444
Db 16423 AACCCCACTACTTTGGGAGGCTGAGCGAGGAAATTCCTTGAAGCCGGGAGGTAGAGTT 16482
QY 445 GCAGTAGCCGAGATAGAGTCACTGCACCTCCAGCCTGGGTGAGCAGAGCAAGACTCCCTC 504
Db 16483 GCGTAAGCCAGATCA-TGCCACTGCACCTCCAGCCTGGCGCAGCAGCAAGACTCTGC 16541
QY 505 TCAGAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 552
Db 16542 TCAAAACAAAAAAGTTGCTATACATATTCAAAAACAATCAATAA 16589

RESULT 7
US-09-764-868-1349/c
; Sequence 1349, Application US/09764868
; Patent No. US2002016871A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ32
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1349
; LENGTH: 24533
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-868-1349

Query Match 29.7%; Score 297.2; DB 9; Length 24533;
Best Local Similarity 80.0%; Pred. No. 6.9e-46;
Matches 399; Conservative 0; Mismatches 94; Indels 6; Gaps 4;
QY 56 TGTAAAAAATAAATGTAATAGGTAGGCTAGCGGTGGCTCAGCCTGTATCCAGCACT 125
Db 6364 TGTCTCAAAATAAATAAATAAGCCAGCGGTGGTGGCGGTGCTGTATCCAGCACT 6305
QY 126 TTAGAAGGTCGAAGAGGGTGGATCACTTGAGGTCAGGAGTTTGAACCCAGCTGGCCAA 185
Db 6304 TTGAGAGCGCGAGGCGGTGGATCACTGAGGTGGGAG-TTAGAGACCAAGCCTGACCA 6246
QY 186 CAGCGTGAACCCCATCTCTACTATAAATAAATAAATAAATAAATAAATAAATAAATAA 245
Db 6245 CATGGAGAAACCCCGTCTCTACTATAAATAAATAAATAAATAAATAAATAAATAAATAA 6186
QY 246 CTGTAATCCAGCACTTTGGGAGGCTGAGACGGGTGGATCAGCTGAAGTCAGGAGTTCAA 305
Db 6187 CTGTAATCCAGCACTTTGAGAGCGCGGCGGTGGATCACTCAGCTCGGAGTTAGA 6128
QY 306 GGCCAGCCTGGCCACATGTTGAACCAAGCTCTCTACTATAAATAAATAAATAAATAAATAA 365
Db 6127 GACCAGCCTGACCAACATGGAGAAACCCCGTCTCTACTATAAATAAATAAATAAATAAATAA 6070
QY 366 TGTGTGGCACACGCTGTAGTCCAGCTACTTTGGGAGGCTGAGCGGCAAGATTCGCTTG 425
Db 6069 CGTGTGGCGCGCTGTATCCAGCTACTCCAGAGGCTGAGGCGATGATTCGCTTG 6010
QY 426 AACCCAGTAGCAGAGGTTGAGTGCAGGAGATGAAGTCACTGCACCTCCAGCCTGGGT 485
Db 6009 AACCCGGGAGCGGAGGTTGCACTGAGCGGAGAT-TCCGCCATTCGACTCCAGCCTAGGC 5951
QY 486 GACAGCAAGACTCCCTCTCAGAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 545
Db 5950 AAGAGTGAGACTCCGTTTCAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5891


```

; LENGTH: 99014
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U78027
US-09-880-107-3428

Query Match      29.4%; Score 293.8; DB 10; Length 99014;
Best Local Similarity 83.8%; Pred. No. 3.3e-45;
Matches 367; Conservative 0; Mismatches 68; Indels 3; Gaps 3;

QY 87 AGGCTAGGCGGTGGCTCAGCGCTGTAAATCCAGCAGCTTTAGAAAGTCGAGAGGGTGG 146
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68994 AGGCTGGCTGGTGGTGTATGCTCTTAATCCAGCAGCTTTGGAGGCGGAGGTGGCGG 68935
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 147 ATCACTTAGGCTCAGGAGTTTGGAGCCAGCCTGGCCAAACACAGGTGAAACCCCATCTCTA 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68934 ATCACTTAGGCGCAGGAG-TTCGAGATCAGCCTGGTCAACATGGCGAATGCCATCTCTA 68876
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 207 CTAAAAATAAAATTAAGCTNGGTGGTGGCTCACACCTGTAAATCCAGCAGCTTTGGG 266
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68875 CTAAAAATAAAATTAAGCTNGGTGGTGGCTCACACCTGTAAATCCAGCAGCTTTGGG 266
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 267 AGGCTGAGAGGGTGGATCACTGAGTCAGGAGTCAAGGCCAGCCTGGCCAGCATGGT 326
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68816 AGGCGAGGTGGTGGATCACTGAGTCAGGAGTCAAGGCCAGCCTGGCCAGCATGGT 68757
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 327 GAAACACAGCTCTCTACTAAAAATACAAAAATTAGCCAGGTGTGGTGGCCACAGCTGTAG 386
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68756 GAAATGCTGCTCTACTAAAAATGCAAAAAATTAAGCTGGCAGCTGGCATGTGCTGTAG 68697
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 387 TCCAGCTACTTGGAGGCTGAGCGGGAAGAAATCGCTTGAACCCAGTAGGAGAGGTTCG 446
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68696 TCCAGCTACTCGGAGGCTAAGCCAGGAGAAATCACTTGAACCCAGTAGGAGAGGTTCG 68637
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 447 AGTGAGCGAGATAAGTCACTGCACTCCAGCCTGGTGGTGGCCAGCAGAGCTCCCTCTC 506
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68636 AGTGAGCGAAAT-CGTGCCACTGCACTCCAGCCTGGTGGTGGCCAGCAGAGCTGTGTCAC 68578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 507 AGAAAAATAAAATAA 524
    ||| ||| ||| |||
Db 68577 CAAAAAATAAAAAA 68560
    ||| ||| ||| |||

RESULT 11
US-09-747-810-1
; Sequence 1, Application US/09747810
; Patent No. US20020012903A1
; GENERAL INFORMATION:
; APPLICANT: Goldman, Steven A.
; APPLICANT: Okano, Hideyuki
; TITLE OF INVENTION: A METHOD FOR ISOLATING AND PURIFYING MULTIPOTENTIAL NEURAL PROGEN
; TITLE OF INVENTION: CELLS AND MULTIPOTENTIAL NEURAL PROGENITOR CELLS
; FILE REFERENCE: 19603/3580
; CURRENT APPLICATION NUMBER: US/09/747,810
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,003
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 52216
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-747-810-1

Query Match      29.2%; Score 291.6; DB 10; Length 52216;
Best Local Similarity 81.8%; Pred. No. 7.7e-45;
Matches 372; Conservative 0; Mismatches 80; Indels 3; Gaps 3;

QY 74 AATTAATGAATAGGCTAGGCGGTGGCTCAGCCTGTAAATCCAGCAGCTTTAGAAGG 133
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 35542 AAAAAATGAATAGGCTGGGCACAGTGTCTCATGCTGTAAATCCAGCTTTGGGAGG 35601
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

QY 134 TCAGAGAGGGTGGATCACTTGAGCTCAGGAGCTTTTGACAGACAGCCTGCCAACACGGTGA 193
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 35602 CTGAGGTGGTGGATCACTTGAGGCGAGGAG-TTCAGGCCACAGCTGGCCAATACGGTGA 35660
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 194 AACCCCATCTCTACTAAAAATA-AAAAATTAGCTNGGTGGTGGCTGCACACTCTTAAT 252
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 35661 AGCCAGCTCTCTCTCTAAAAATACAAAAATTAGCCGGAGCAGTAGTGGCACGCACTGCAAT 35720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 253 CCAGCAGCTTTGGGAGGCTGAGACGGGTGGATCACCCTGAAGTCAGGAGTTCAAGGCCAGC 312
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 35721 CCAGCAGCTTTGGGAGGCTGAGGCGGGAATCACCCTGAGGTGAGGATTCGAGGCCAGC 35780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 313 CTGGGCAACATCGTTCGAAACACGCTCTCTACTAAAAATACAAAAATAGCAGAGTGTGTG 372
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 35781 CTGGGCAACATCGTTCGAAACACGCTCTCTACTAAAAATACAAAAATAGTGGCATGTG 35840
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 373 GCACAGCCTGTAGTCCAGCTACTTTGGGAGGCTGAGGCGGGAAGAAATCGTTGAACCCAG 432
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 35841 GCGTGTGCTCTAATCCAGCTACTTTGGGAGGCTGGGCGAGGAGAAATCTCTTGAACCCAG 35900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 433 TAGGCAGAGGTTCAGTTCGAGCGGAGATAAGAGTCACTGCACTCCAGCCTGGGTGACAGAG 492
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 35901 GAGGCAGAGGTTCAGTTCGAGCGGAGAT-TGCACCACTGCGCTCCACACTGGGCCACAGAG 35959
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 493 CRAAGACTCCCTCTCAGAAAAATAAAATAAAAAA 527
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 35960 TGAGACTCCATCTCAAAAAAATAAAAAA 35994
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
US-09-818-656A-3/c
; Sequence 3, Application US/09818656A
; Patent No. US20020142381A1
; GENERAL INFORMATION:
; APPLICANT: GONG, Pangcheng et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; FILE REFERENCE: AND USES THEREOF
; FILE REFERENCE: CL001191
; CURRENT APPLICATION NUMBER: US/09/818,656A
; CURRENT FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 40645
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-818-656A-3

Query Match      29.1%; Score 291.2; DB 10; Length 40645;
Best Local Similarity 80.6%; Pred. No. 8.9e-45;
Matches 402; Conservative 0; Mismatches 89; Indels 8; Gaps 5;

QY 69 AAAAAAATAAATTAAGTAAGCTAGGCTAGGCGGTGGCTGCAGGCTGTATATCCAGCACTTTA 128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16379 AAAAAAAGAAAGAAAGAAAGCGAGGTGTGTGGATTACCGCTGTATATCCAGCACTGTG 16320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 129 GAAGTCCGAGAGGGTGGATCACTTGAGTTCAGGAGTTTGGAGACCAAGCTGGCCAAACAC 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16319 GGAGGCCAAGCGCAGATAC--GAGGTCAAGAG-ATCAGAGACCATCTCTGGCCAAAT 16263
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 189 GGTGAACCCCATCTCTACTAAAAATAAAAAATAGCTNGGTGGTGGCTGCACACCTG 248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16262 GGTGAACCCCATCTCTAGTAAAAATAAAAAATAGCTGGGCGGTGGCTCAAGGCTG 16203
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 249 TAAATCCAGCAGCTTTGGGAGGCTGAGAGGGTGGATCACTTGAAGTCAGGAGTTCAAGGC 308
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16202 TAAATCCAGCAGCTTTGGGAGGCTGAGGCGGTGGATCA--TGAGGTTCAGGAGATCGAGC 16145
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 309 CAGCCTGGCAACATGGTGAACACAGCTCTCTACTAAAAATACAAAAATTAAGGAGT 366
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16144 CATCTGGCTAACACAGTGAACCCCATCTCTACTAAAAATAAAAAATTAAGGAGT 16085
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

QY 367 GTGGTGGACACGCCCTGTAGTCCAGCTACTTGGGAGGCTGAGCGGAGAAATCGCTTGA 426
Db 16084 GTGGTGGTGGTGGCTTATAGTCCAGCTACTTGGGAGGCTGAGGAGAGAAATGCGGTGA 16025
QY 427 ACCAGTAGGACAGAGGTTCAGTGGAGCGGAGATAGAGTCACTCCAGCTCGACCTGGGTG 486
Db 16024 ACCAGTAGGCGGAGGTTCAGTGGAGCGGAGAT-TGTGCCACTCAATCCAGCTGGGTG 15966
QY 487 ACAGAGCAAGACTCCCTCTCAGAAAAATAAATAAATAAATAAATAAATAAATAAATAA 546
Db 15965 ACAGAGCAAGACTCCCTCTCAAAAAATAAAGTAAATAAATAAATAAATAAATAAATA 15906
QY 547 ATAAATAAATAATCTAAAA 565
Db 15905 AATAAATAAATAAATAAATAA 15887
RESULT 13
US-09-933-267A-1/c
; Sequence 1, Application US/09933267A
; Patent No. US20020123095A1
; GENERAL INFORMATION:
; APPLICANT: Kalush, Francis et al.
; TITLE OF INVENTION: Estrogen receptor alpha variants and
; FILE OF INVENTION: methods of detection thereof
; FILE REFERENCE: CLO00258C14
; CURRENT APPLICATION NUMBER: US/09/933,267A
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 60/160626
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: 60/183756
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 09/692414
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/768184
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 09/804076
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 09/826314
; PRIOR FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 465237
; TYPE: DNA
; ORGANISM: human
US-09-933-267A-1

Query Match 29.1%; Score 290.8; DB 10; Length 465237;
Best Local Similarity 81.3%; Pred. No. 1.3e-44;
Matches 373; Conservative 0; Mismatches 83; Indels 3; Gaps 3;
QY 78 AAATGGAATAGGTAGCGGGTGGCTCAGCGCTGTATCCAGCACTTTAGAGGTGCA 137
Db 35056 AAAAGGTTTAGGGGGGCGATGGTGGCTCAGCGCTGTATCCAGCACTCTCGGAGGCCGA 34997
QY 138 AGAGGTGATCACTTGAGTTCAGAGTTCAGACACAGCGCTGGCCACACAGGTGAACCC 197
Db 34996 GGTGGTGGATCACTGAGGTTCAGAGT-TCAAGACAGGATGGCCACATGTTGAACCC 34938
QY 198 CCATCTCTACTAAATAAATA-AAAAATTAGTNGGTGGGTGGCTCACACCTGTAATCCCA 256
Db 34937 TCGTCTCTATTAAAAATACAAAAAATGACCTGGCGAGTGGCTCACCGCTGTATCCCA 34878
QY 257 GCATTTGGGAGGTGAGACGGGTGGATCACTGAAGTCAGGAGTTCAAGCCAGCGCTGG 316
Db 34877 GCATTTGGGAGCGGAGGAGGTGGATCACTGAGTTCAGAGTTTGAACAGCGCTGG 34818
QY 317 GCACATGTTGAACACCGCTCTCTACTAAAAATAACAAAAATTAGCCAGGTGTGGTGCAC 376
Db 34817 CCAACATGCGGAACCCCGACTCTCTAACAATAACAAAAATTAGTGGCAGCGTGGAT 34758
QY 377 ACCGCTGTAGTCCCGAGTACTTGGGAGGCTGAGCGCGGAGAAATCGCTTGAACCCAGTAGG 436

Db 34757 GTGCTGTATCCAGCTACTAGGAGGCTGAGCGAGGAAATCGCTGGAACCTGGAGG 34698
QY 437 CAGAGTTCAGTGGAGCGGAGATAGAGTCACTCCAGCTCGGCTGGGTGACAGAGCAAG 496
Db 34697 CAGAGTTCAGTGGAGCGGAGAT-GGTGCCACTGCTCCAGCTGGGCAACAGAGCAAG 34639
QY 497 ACTCCCTCTCAGAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 535
Db 34638 ACTCCCTCTCAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 34600

RESULT 14
US-09-764-878-363
; Sequence 363, Application US/09764878
; Patent No. US20020090615A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAL21
; CURRENT APPLICATION NUMBER: US/09/764,878
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 428
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 363
; LENGTH: 3273
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-878-363

Query Match 29.0%; Score 289.8; DB 10; Length 3273;
Best Local Similarity 84.5%; Pred. No. 1.3e-44;
Matches 372; Conservative 0; Mismatches 63; Indels 5; Gaps 4;
QY 88 GGCTAGGCGGTGGTTCAGCGCTGTATCCAGCACTTTAGAGGTTCGAGAGGTGGA 147
Db 2686 GCGCAGGCGCGGTAGTTCAGCGCTGTATCCAGCACTTTGGAGGCTGAGAGGCGGA 2745
QY 148 TCACITGAGTTCAGGAGTTCAGACAGCGCTGGCCACAGCGTGAACCCCATCTCTAC 207
Db 2746 TCACITGAGTTCGCGGAGTTT-GAGACAGCGCTGGCCACATGTTGAACCTCTCTAT 2804
QY 208 TAAAAATAAATAATAGCTNGGTGGTGGCTCAGCTGTATCCAGCACTTTGGAGGTGGA 267
Db 2805 TAAAAATAAATAATAGCGCGCGGTGGCTCAAGCTGTATATCCAGCACTTTGGGA 2864
QY 268 GGCTGAGACGGGTGGATCAGCTGAAGTCAGGAGTTCAAGGCGCAGCTGGGCAACATGTG 327
Db 2865 GCGCAGGCGCGCGGATCAC--GAGGTACAGGATCGAGACCTTCCTGGCTAACAGGTG 2922
QY 328 AAACCACTCTCTACTAAAAATAC-AAAAATTAGCAGGTGTGGTGGCAGCGCTGTAG 386
Db 2923 AAACCCATCTCTACTAAAAATACAAAAATTAGCGGCGGTGGTGGCGGCGCTGTAA 2982
QY 387 TCACAGCTTCTGGGAGGTGAGCGGAGAAATCGCTTGAACCCAGTTCAGGAGGTTCG 446
Db 2983 TCACAGCTTCTGGGAGGTGAGCGGAGGATGGCGTGAACCCAGGAGGCGGAGTTCG 3042
QY 447 AGTGGCCGAGATAGAGTCACTGCATCCAGCGTGGGTGACAGAGAGTCCCTCTC 506
Db 3043 AGTGGCCGAGAT-AGCGCCACTGCATCCAGCGTGGCGGAGAGTGGAGTCTCCCTCTC 3101
QY 507 AGAAAAATAAATAAATAA 526
Db 3102 CAAAAAATAAATAAATAA 3121

RESULT 15
US-09-764-878-362
; Sequence 362, Application US/09764878
; Patent No. US20020090615A1
; GENERAL INFORMATION:

Search completed: November 24, 2002, 10:38:37
Job time : 930 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2002, 07:40:52 ; Search time 71 Seconds
(without alignments)
4319.391 Million cell updates/sec

Title: US-09-784-423-32
Perfect score: 1000
Sequence: 1 GGTGGACCTTACCTCTCT.....CAGATGATGACGGCGTGC 1000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	999	99.9	1000	US-09-018-584A-32	Sequence 32, Appl
2	296	29.6	15297	US-09-817-180-3	Sequence 3, Appl
3	284	28.4	2713	US-08-916-901-6	Sequence 6, Appl
4	284	28.4	2713	US-09-154-602-6	Sequence 6, Appl
5	284	28.4	7720	US-09-318-448-5	Sequence 5, Appl
6	277.8	27.8	7210	US-08-257-963B-10	Sequence 10, Appl
7	277.8	27.8	7210	US-08-367-841A-10	Sequence 10, Appl
8	277.8	27.8	7210	US-08-367-841A-10	Sequence 10, Appl
9	277.8	27.8	14581	PCT-US95-07201-10	Sequence 10, Appl
10	277.8	27.8	22481	US-08-520-373D-4	Sequence 4, Appl
11	277.8	27.8	22481	US-08-367-841A-43	Sequence 43, Appl
12	277.8	27.8	22481	PCT-US95-07201-43	Sequence 43, Appl
13	275.6	27.6	112132	US-09-875-223-2	Sequence 2, Appl
14	272.6	27.3	29629	US-09-729-995-3	Sequence 3, Appl
15	271.4	27.1	62804	US-09-800-960-3	Sequence 3, Appl
16	269.6	27.0	59065	US-09-813-817-3	Sequence 3, Appl
17	269.6	27.0	59065	US-09-978-197-3	Sequence 3, Appl
18	267.4	26.7	14581	US-08-520-373D-4	Sequence 4, Appl
19	266.6	26.7	17327	US-07-906-871-15	Sequence 15, Appl
20	265.8	26.6	7210	US-08-257-963B-10	Sequence 10, Appl
21	265.8	26.6	7210	US-08-367-841A-10	Sequence 10, Appl
22	265.8	26.6	7210	US-08-367-841A-10	Sequence 10, Appl
23	265.8	26.6	22481	PCT-US95-07201-10	Sequence 10, Appl
24	265.8	26.6	22481	US-08-367-841A-43	Sequence 43, Appl
25	265.8	26.6	22481	PCT-US95-07201-43	Sequence 43, Appl
26	263.8	26.4	3267	US-09-875-223-2	Sequence 2, Appl
27	263.8	26.4	3267	US-08-257-963B-12	Sequence 12, Appl
				US-08-367-841A-12	Sequence 12, Appl

Sequence 12, Appl
Sequence 5, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 1, Appl
Sequence 20, Appl
Sequence 21, Appl
Sequence 22, Appl
Sequence 3, Appl
Sequence 21, Appl
Sequence 21, Appl
Sequence 2, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 7, Appl
Sequence 3, Appl
Sequence 3, Appl

c 28 263.8 26.4 3267 5 PCT-US95-07201-12
c 29 263.8 26.4 5282 4 US-08-520-373D-5
c 30 262.8 26.3 10380 4 US-09-077-354B-3
c 31 261 26.1 16389 4 US-09-741-154-3
c 32 258.6 25.9 162450 4 US-09-345-882-1
c 33 258.6 25.9 246240 2 US-08-724-394A-20
c 34 258.6 25.9 246240 2 US-08-724-394A-21
c 35 258.6 25.9 246240 2 US-08-724-394A-22
c 36 257 25.7 10684 3 US-08-618-100B-3
c 37 256.6 25.7 1363 1 US-08-776-088-21
c 38 256.6 25.7 1363 5 PCT-US95-09145A-21
c 39 256.2 25.6 53526 3 US-08-658-136-2
c 40 256.2 25.6 53577 3 US-08-658-136-1
c 41 253.6 25.4 4192 4 US-09-122-128B-1
c 42 253.2 25.3 162450 4 US-09-345-882-1
c 43 249.4 24.9 11811 4 US-09-078-294-7
c 44 248.4 24.8 59065 4 US-09-813-817-3
c 45 248.4 24.8 59065 4 US-09-978-197-3

ALIGNMENTS

RESULT 1
US-09-018-584A-32
; Sequence 32, Application US/09018584A
; Patent No. 6238863
; GENERAL INFORMATION:
; APPLICANT: Schumm, James W.
; TITLE OF INVENTION: MATERIALS AND METHODS FOR IDENTIFYING AND ANALYZING INTERMEDIATE TANDEM REPEAT DNA MARKERS
; TITLE OF INVENTION: REPEAT DNA MARKERS
; NUMBER OF SEQUENCES: 147
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Promega Corporation
; STREET: 2800 Woods Hollow Road
; CITY: Madison
; STATE: Wisconsin
; COUNTRY: U.S.A.
; ZIP: 53711-5399
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Mb
; COMPUTER: IBM compatible PC
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Word 97 (DOS text format)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/018,584A
; FILING DATE: 04-Feb-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Grady J. Frenchick
; REGISTRATION NUMBER: 29,018
; REFERENCE/DOCKET NUMBER: 16026.9180
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 257-3501
; TELEFAX: (608) 257-2275
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1000 bp
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Circular
; MOLECULE TYPE: Genomic DNA
; HYPOTHETICAL: no
; IMMEDIATE SOURCE:
; CLONE: SI32
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 22
US-09-018-584A-32

Query Match 99.9%; Score 999; DB 4; Length 1000;
Best Local Similarity 100.0%; Pred. No. 2.4e-217;

Matches 1000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GETGTGACCTTATCTCTGAACTCAGTTTCTCTATCCGTAAATGAAGCTGCTAG 60
Db 1 GETGTGACCTTATCTCTGAACTCAGTTTCTCTATCCGTAAATGAAGCTGCTAG 60

QY 61 ATGTGTTGAAAAAATAATGAATAGGCTAGGCGGGTGGCTCAGCTGTATATCCCA 120
Db 61 ATGTGTTGAAAAAATAATGAATAGGCTAGGCGGGTGGCTCAGCTGTATATCCCA 120

QY 121 GCACCTTTAGAGGTGCAAGAGGGTGGATCAGTGGGCTGAGGCTTTTGGAGCAGGCTG 180
Db 121 GCACCTTTAGAGGTGCAAGAGGGTGGATCAGTGGGCTGAGGCTTTTGGAGCAGGCTG 180

QY 181 GCCAACCGGTGAAACCCCATCTCTACTAAAAATAAAAAATAGCTNGGGTGGCT 240
Db 181 GCCAACCGGTGAAACCCCATCTCTACTAAAAATAAAAAATAGCTNGGGTGGCT 240

QY 241 CACACCTGTATATCCAGCAGCTTTGGGAGGTGAGACGGGTGGATCAGCTGAAGTCAGGAG 300
Db 241 CACACCTGTATATCCAGCAGCTTTGGGAGGTGAGACGGGTGGATCAGCTGAAGTCAGGAG 300

QY 301 TTCAGGCGAGCTGGGCAACATGGTGAACACAGCTCTCTACTAAAAATAAAAAATAG 360
Db 301 TTCAGGCGAGCTGGGCAACATGGTGAACACAGCTCTCTACTAAAAATAAAAAATAG 360

QY 361 CCAGGTGTGTGGCACACGCTGTAGTCCAGCTACTTTGGGAGGCTGAGGCGGAGATC 420
Db 361 CCAGGTGTGTGGCACACGCTGTAGTCCAGCTACTTTGGGAGGCTGAGGCGGAGATC 420

QY 421 GCTTGAACCCAGTAGGAGAGGTGAGTGGAGAGTGAAGTACCTGACCTCCAGCC 480
Db 421 GCTTGAACCCAGTAGGAGAGGTGAGTGGAGAGTGAAGTACCTGACCTCCAGCC 480

QY 481 TGGGTGACAGAGCAAGACTCCCTCTCAGAAATAAATAAATAAATAAATAAATAA 540
Db 481 TGGGTGACAGAGCAAGACTCCCTCTCAGAAATAAATAAATAAATAAATAAATAA 540

QY 541 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 600
Db 541 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 600

QY 601 AGTAATAGTATCAATATCCCAACCTTACACACTGCTGCTGAATTTAGTTTCTTTTGT 660
Db 601 AGTAATAGTATCAATATCCCAACCTTACACACTGCTGCTGAATTTAGTTTCTTTTGT 660

QY 661 ACCCCCATTAGACTTAAGGAGAGTCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 661 ACCCCCATTAGACTTAAGGAGAGTCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCT 720

QY 721 GGCACATAGTTGGGTCTCAGTGAACATGTTGAGTGAATGAGCAATGCAAGAACTCC 780
Db 721 GGCACATAGTTGGGTCTCAGTGAACATGTTGAGTGAATGAGCAATGCAAGAACTCC 780

QY 781 AGGCATCTGGGAGCCCTCCAGGCGGTGAGTTCGGGAACTCATAGTCTGTCTCAAT 840
Db 781 AGGCATCTGGGAGCCCTCCAGGCGGTGAGTTCGGGAACTCATAGTCTGTCTCAAT 840

QY 841 GGCCCACTGAAGGTAGAGAGTTCTGGGTCCCACTCCGACCCCACTCTCTGACTAC 900
Db 841 GGCCCACTGAAGGTAGAGAGTTCTGGGTCCCACTCCGACCCCACTCTCTGACTAC 900

QY 901 TGCTGAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 960
Db 901 TGCTGAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 960

QY 961 CAGGACTGCAAGGAGCCAGCAGATGATGACGGGTGC 1000
Db 961 CAGGACTGCAAGGAGCCAGCAGATGATGACGGGTGC 1000

RESULT 2

US-09-817-180-3

; Sequence 3, Application US/09817180
; Patent No. 6340584
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: THEREOF
; CURRENT APPLICATION NUMBER: US/09/817.180
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 15297
; TYPE: DNA
; ORGANISM: Human
US-09-817-180-3

Query Match 29.6%; Score 296; DB 4; Length 15297;
Best Local Similarity 83.4%; Pred. No. 3.1e-58;
Matches 371; Conservative 0; Mismatches 71; Indels 3; Gaps 3;

QY 97 CGTGTGCTCAGCGCTGTATCCAGCAGCTTTAGAGGTCGAAGGGTGGATCAGTGGAG 156
Db 11700 CAGTTGCTCAGCGCTGTATCCAGCAGCTTTGGGAGGCTGAGCTGGGTGATCAGTTGAG 11759

QY 157 GTCAGGAGTTTGGAGCAGCGCTGCCAACACAGCTGAACCCCATCTCTACTAAAAATAA 216
Db 11760 CCCAGGAG-TTCAAGATCAGCTTGGACACACAGTGAACCTTCTGTAAAAATAATAC 11818

QY 217 AAATAGCTTNGGTGGGTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAGAC 276
Db 11819 AAAAAATAGACTGGGCGCGGTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCGGAGC 11878

QY 277 GGGTGTGATCACTGAACTCAGGAGTCAAGCCAGCGCTGGGCAACATGTTGAAACCCCAT 336
Db 11879 AGCTGGATCAGCTGTGGTCAGGAGTTGAGACCAAGCAGCAACATGTTGAAACCCCAT 11938

QY 337 CTCTACTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 396
Db 11939 CTCTACTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 11998

QY 397 TTGGGAGGCTCAGCGGAGAGATCGTTGAACCCAGTAGGACAGAGTTCAGTGGAGCGGA 456
Db 11999 TTGGGAGGCTGAGGTGGGAGATTCGTTGAACCCAGAGGCGGAGGCTGAGTGGAGCGGA 12058

QY 457 GATAGAGTCACTGCACTCCAGCTCCAGCTGGGTGGTGC-AGAGCAAGACTCCCTCTCAGAAAAATAA 515
Db 12059 GAT-TGTCGCACTGCACTCCAGCTCCAGCTGGGTGGTGGGAGCAAGAGTGAACCTCTCAAAAAATAA 12117

QY 516 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 540
Db 12118 CAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 12142

RESULT 3

US-08-916-901-6
; Sequence 6, Application US/08916901
; Patent No. 5892012
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: RAB PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/916,901
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0367 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2713 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-916-901-6

Query Match 28.4%; Score 284; DB 2; Length 2713;
Best Local Similarity 81.3%; Pred. No. 1.1e-55;
Matches 365; Conservative 0; Mismatches 81; Indels 3; Gaps 3;
QY 74 AATTAATGGAATAGCTAGCGCGGTGGCTCAGCGCTGTAATCCAGCATTAGAAGG 133
Db 1688 AAAAAACAAAATAGCGCGGTGGCTCAGCGCTGTAATCCAGCATTAGAAGG 1747
QY 134 TCGAAGAGGTGGATCAGCTTGGAGTCAAGAGTTTGGAGCAGCGCTGCGCAACACGGTGA 193
Db 1748 CTGAGGAGGAGATCAGCTTGGAGTCAAGAG-TTTGAGACAGCGCTGGCAACGGTGA 1806
QY 194 ACCCCATCTCTACTAAAAATAAAAAATAGCTNGGGTGGCTGCGCTCAGCAGCTGTAATC 253
Db 1807 AACCTGTCTCTACTAAAAATAAAAAATAGCTNGGGTGGCTGCGCTCAGCAGCTGTAATC 1866
QY 254 CCAGCAGCTTTGGAGGCTGAGACGGGTGGATCACTGAGTCAAGAGTTTCAAGGCCAGCC 313
Db 1867 TCAGCAGCTTTGGAGGCGGAGCAGGTGGATCACTGAGTCAAGAGTTTCAAGGCCAGCC 1926
QY 314 TGGCAACATGGTGAACACAGCTCTCTACTAAAAATAAAAAATAGCAGGTGGTGG 373
Db 1927 TGGCAACATGGTGAACACAGCTCTCTACTAAAAATAAAAAATAGCAGGTGGTGG 1986
QY 374 CACAGCGCTGTAGTCCAGCTACTTGGAGGCTGAGCGGAGAGTCAAGTCAAGCCAGT 433
Db 1987 TGCATGCTGTATCCAGCTACTTGGAGGCTGAGCGGAGAGTCAAGTCAAGCCAGT 2046
QY 434 AGGAGAGGTTCAGTCAAGCGAGATAGAGTCACTGCACTCCAGCTGGGTGACAGAGC 493
Db 2047 A-GCAGAGGTTCAGTCAAGCGAGATCA-TGCCACTGCACTCCAGCTGGGTGACAGAGC 2104
QY 494 AAGACTCCCTCTCAGAAAAATAAAAA 522
Db 2105 GAGACTCCATTTCAAAAAAAGAACTACA 2133

RESULT 4
US-09-154-602-6
Sequence 6, Application US/09154602
Patent No. 6300472
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: RAB PROTEINS

NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/154,602
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/916,901
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0367 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2713 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-154-602-6

Query Match 28.4%; Score 284; DB 4; Length 2713;
Best Local Similarity 81.3%; Pred. No. 1.1e-55;
Matches 365; Conservative 0; Mismatches 81; Indels 3; Gaps 3;
QY 74 AATTAATGGAATAGCTAGCGCGGTGGCTCAGCGCTGTAATCCAGCATTAGAAGG 133
Db 1688 AAAAAACAAAATAGCGCGGTGGCTCAGCGCTGTAATCCAGCATTAGAAGG 1747
QY 134 TCGAAGAGGTGGATCAGCTTGGAGTCAAGAGTTTGGAGCAGCGCTGCGCAACACGGTGA 193
Db 1748 CTGAGGAGGAGATCAGCTTGGAGTCAAGAG-TTTGAGACAGCGCTGGCAACGGTGA 1806
QY 194 ACCCCATCTCTACTAAAAATAAAAAATAGCTNGGGTGGCTGCGCTCAGCAGCTGTAATC 253
Db 1807 AACCTGTCTCTACTAAAAATAAAAAATAGCTNGGGTGGCTGCGCTCAGCAGCTGTAATC 1866
QY 254 CCAGCAGCTTTGGAGGCTGAGACGGGTGGATCACTGAGTCAAGAGTTTCAAGGCCAGCC 313
Db 1867 TCAGCAGCTTTGGAGGCGGAGCAGGTGGATCACTGAGTCAAGAGTTTCAAGGCCAGCC 1926
QY 314 TGGCAACATGGTGAACACAGCTCTCTACTAAAAATAAAAAATAGCAGGTGGTGG 373
Db 1927 TGGCAACATGGTGAACACAGCTCTCTACTAAAAATAAAAAATAGCAGGTGGTGG 1986
QY 374 CACAGCGCTGTAGTCCAGCTACTTGGAGGCTGAGCGGAGAGTCAAGTCAAGCCAGT 433
Db 1987 TGCATGCTGTATCCAGCTACTTGGAGGCTGAGCGGAGAGTCAAGTCAAGCCAGT 2046
QY 434 AGGAGAGGTTCAGTCAAGCGAGATAGAGTCACTGCACTCCAGCTGGGTGACAGAGC 493
Db 2047 A-GCAGAGGTTCAGTCAAGCGAGATCA-TGCCACTGCACTCCAGCTGGGTGACAGAGC 2104
QY 494 AAGACTCCCTCTCAGAAAAATAAAAA 522
Db 2105 GAGACTCCATTTCAAAAAAAGAACTACA 2133

RESULT 5
US-09-318-448-5/c

```
; Sequence 5, Application US/09318448
; Patent No. 6210950
; GENERAL INFORMATION:
; APPLICANT: Johnson, William G.
; APPLICANT: Stearns, Edward S.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; TITLE OF INVENTION: DEVELOPMENTAL DISORDERS
; FILE REFERENCE: 601-1-057
; CURRENT APPLICATION NUMBER: US/09/318,448
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7720
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-318-448-5

Query Match      28.4%; Score 284; DB 4; Length 7720;
Best Local Similarity 84.9%; Pred. No. 1.4e-55;
Matches 376; Conservative 0; Mismatches 61; Indels 6; Gaps 5;

QY 87 AGCTAGCGCGGTGGCTCAGCCTGTATCCAGCACTTTAGAGGTGGAAGAGGGTGG 146
Db 6233 AGCTGGGTGCGAGTGGCTTACTTGTATATCCAGCACTTTGGAGGCCAAGCGAGTGG 6174
QY 147 ATCACTTGAGTCAGAGTTTGAGACAGCCTGGCCACAGCGGTGAACCCCATCTCTA 206
Db 6173 ATCACTTGAGTCAAGAG--TTTGAGACCAAGCCTGGCCAAATGTGAACCCCGTCTCTA 6115
QY 207 CTAAATAAATA-AAAAATTAGCTNGGTCGGTGGCTCACACCTGTATATCCAGCACTTTGG 265
Db 6114 CTAAATAAATAAATAATAGCGCGGTGGCTGGCTCAGCCTGTATATCCAGCACTTTGG 6055
QY 266 GAGGCTGAGCGGGTGATCACCTGAAGTCAGAGTTCAAGGCCAGCCTGGGCAACATGG 325
Db 6054 GAGCGCCAGCGGGTGGATCAC--GAGGCTAAGAGATCAAGACCATCTCTGGCCACATGG 5997
QY 326 TGAACACCACTCTCTACTAAAAATAC-AAAAATTAGCAGGTGTGGTGGCAGACGCTGT 384
Db 5996 TGAACCCCTCTCTACTAAAAATACAAAAAATTAGCGGCGGTGGTGGCTGGCACCTGT 5937
QY 385 AGTCCAGCTACTTGGGAGCTGAGCGGGAAGATCGCTTGAACCCAGTAGCGAGAGTT 444
Db 5936 AGTCCAGCTACTTGGGAGCTGAGCGGAGAGATGTGTGAACCCGCGAGGTGAGCTT 5877
QY 445 GCAGTGAGCGCGAGATAGAGTCAGTCACTCCAGCCTGGGTGGTACAGAGCAAGACTCCCTC 504
Db 5876 GCAGTGAGCGGAGATCA-TGCCACTGCACTCCAGCCTGGGCGAGACAGCAAGACTCCATC 5818
QY 505 TCAGAAAAATAAATAAATAAATAA 527
Db 5817 TCAGAAAAATAAACAACAATAA 5795

RESULT 6
US-08-257-963B-10
; Sequence 10, Application US/08257963B
; Patent No. 5840686
; GENERAL INFORMATION:
; APPLICANT: Chader, Gerald J.; Becerra, S.
; APPLICANT: Patricia; Schwartz, Joan P.;
; APPLICANT: Taniwaki, Takayuki
; TITLE OF INVENTION: PIGMENT EPITHELIUM
; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION OF ITS NOVEL
; TITLE OF INVENTION: BIOLOGICAL ACTIVITY AND SEQUENCES ENCODING
; TITLE OF INVENTION: AND EXPRESSING THE PROTEIN
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan
; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
```

```
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/257,963B
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/952,796
; FILING DATE: 24-SEPT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36434
; REFERENCE/DOCKET NUMBER: 20264126US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7210 Base Pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Human
; IMMEDIATE SOURCE:
; LIBRARY: DASH II
; FEATURE:
; NAME/KEY: JT106
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: 7.2 kb No. 5840686 1 fragments
; OTHER INFORMATION: Derived from human placental genomic DNA
US-08-257-963B-10

Query Match      27.8%; Score 277.8; DB 2; Length 7210;
Best Local Similarity 82.3%; Pred. No. 3.5e-54;
Matches 354; Conservative 0; Mismatches 73; Indels 3; Gaps 3;

QY 88 GCGTAGGCGCGGTGGCTCAGCCTGTATATCCAGCACTTTAGAGGTGGAAGAGGGTGG 147
Db 3756 GCGCGGACAGCGTGGCTCAGCCTGTAGTCCAGCACTTTGGAGGCCGAGGCGGCGAGA 3815
QY 148 TCACCTGAGGTCAGGAGTTTGGAGACCAAGCCTGGCCCAACAGCGGTGAACCCCATCTCTAC 207
Db 3816 TCACCTGAGGTCAGGAG-TTCGAGACCAAGCCTGGCTTAACAGATGAACCCCGTCTCTAC 3874
QY 208 TAAAAATA-AAAAATTAGCTNGGTCGGTGGCTCACACCTGTATATCCAGCACTTTGGG 266
Db 3875 TAAAAATACAAAAAATTAGCTGGGCAAGGTGGCTGTGCTGTAAATCCAGCACTTTGGG 3934
QY 267 AGGCTGAGCGGGTGGATCACTCAAGTCAGGAGTTCAAGGCCAGCCTGGGCAACATGGT 326
Db 3935 AGGCGAGAGTGGGCGAGTCACTTCAAGTCAGGAGTTTGGAGACCAAGCCTAGCCACATGGT 3994
QY 327 GAAACCAAGTCTCTACTAAAAATA-AAAAATTAGCCAGGTGTGGTGGCAGACGCTGTAG 386
Db 3995 GAAACCCCATCTCTACTAAAAATA-AAAAATTAGCCAGGTGTGGTGGCAGTGTGTA 4054
QY 387 TCCAGCACTCTGGGAGCTGAGCGGGAAGATCCCTTGAACCCAGTAGGAGAGGTTCG 446
Db 4055 TCCAGCGCAGTACGAGGCTGAGGCGAGGAATCACTGGAATCTCTGGAGTGGAGTGGC 4114
QY 447 AGTGAGCGGAGATAGAGTCACTCCAGCCTGGTGGTGAACAGACAGCAAGACTCCCTCTC 506
Db 4115 AGTGAGCGGAGAT-GGTACCTCTGTACTCCAGCCTGGGCGGAGAGTGAAGACTCCGCTCTC 4173
QY 507 AGAAAAATAA 516
```

Db 4174 AAAAAAAAAA 4183

RESULT 7

US-08-367-841A-10

; Sequence 10, Application US/08367841A

; Patent No. 6319687

; GENERAL INFORMATION:

; APPLICANT: Chader, Gerald J.; Rodriguez,

; APPLICANT: Ignacio R.; Mazuruk, Krzysztof;

; APPLICANT: Tombran-Tink, Joyce

; TITLE OF INVENTION: PIGMENT EPITHELIUM

; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC

; TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PEDF GENE

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Morgan & Finnegan

; STREET: 345 Park Avenue

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy Disk

; COMPUTER: IBM PC Compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WORDPERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/367,841A

; FILING DATE: 30-DEC-1994

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/257,963

; FILING DATE: 07-JUN-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/952,796

; FILING DATE: 24-SEP-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: DOROTHY R. AUTH

; REGISTRATION NUMBER: 36434

; REFERENCE/DOCKET NUMBER: 20264126US2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 758-4800

; TELEFAX: (212) 751-6849

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 7210 Base Pairs

; TYPE: Nucleic Acid

; STRANDEDNESS: Double

; TOPOLOGY: Unknown

; MOLECULE TYPE: Genomic DNA

; ORIGINAL SOURCE:

; ORGANISM: Human

; IMMEDIATE SOURCE:

; LIBRARY: DASH II

; FEATURE:

; NAME/KEY: JT6A

; LOCATION:

; IDENTIFICATION METHOD:

; OTHER INFORMATION: 7.0 Kb No. 6319687 1-No. 6319687

; OTHER INFORMATION: fragment; derived from human placental

; OTHER INFORMATION: genomic DNA; also referred to as JT106

US-08-367-841A-10

Query Match 27.8%; Score 277.8; DB 4; Length 7210;

Best Local Similarity 82.3%; Pred. No. 3 5e-54;

Matches 354; Conservative 0; Mismatches 73; Indels 3; Gaps 3;

QY 88 GCGTAGGCGGCGGCTACCGCTGTATCCAGACATTAGAAAGGTCCGAAGAGGGTGA 147

Db 3756 GCGCGGACCGGTGGCTACGCGCTGTAGTCCAGCATTGGGAGGCCGAGGCGCAGA 3815

QY 148 TCACCTGAGGTCAGGAGCTTTTGGACACCGCTGGGCAACACGGTGAACCCCACTCTAC 207

Db 3816 TCACCTGAGGTCAGGAG-TTCGAGACCGCTGGCTAACACGATGAACCCCGCTCTAC 3874

QY 208 TAAAAATA-AAAAAATTAGCTNGGCTGGCTCACACCTGTAAATCCACACACTTTGG 266

Db 3875 TAAAAATACAAAAATTAGCTGGGACGGTGGCTCGTCTGTATCCACGACTTTGG 3934

QY 267 AGGCTGAGACGGGTGGATCACCTGAAGTCAGGAGTTCAGGGCCAGCTGGGCAACATGGT 326

Db 3935 AGGAGAGGTGGGAGATCATTGAGGTCAGGAGTTTGGAGACCGCTAGCCACATGGT 3994

QY 327 GAAACCACTCTCTACTAATAAATACAAAAATAGCCAGGTGGTGGGCACACGCTGTAG 386

Db 3995 GAAACCCATCTCTACTAATAAATACAAAAATAGCCGAGGTGGTGGCACGCTGTAA 4054

QY 387 TCCAGCTACTTGGGAGGCTGGAGGGGGAAGTGGCTTGAACCCAGTAGGACAGCTTCC 446

Db 4055 TCCAGCCAGTCAGGAGGCTGGAGGAGGAGTACTTGAATCTCTGGAGGTGGAGGTGG 4114

QY 447 AGTGAGCCGAGATAGAGTCACTGCCTCCAGGCTGGGTGACAGAGCAAGACTCCCTCTC 506

Db 4115 AGTGAGCCGAGAT-GGTACCTCTGTACTCCAGGCTGGGAGGAGTGGAGTCCGCTCTC 4173

QY 507 AGAAATAAA 516

Db 4174 AAAAAAAAAA 4183

RESULT 8

PCT-US95-07201-10

; Sequence 10, Application PC/TUS9507201

; GENERAL INFORMATION:

; APPLICANT: Chader, Gerald J.; Becerra, Sofia

; APPLICANT: Patricia; Schwartz, Joan P.;

; APPLICANT: Taniwaki, Takayuki

; TITLE OF INVENTION: PIGMENT EPITHELIUM

; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC

; TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PEDF GENE

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Morgan & Finnegan, L.L.P.

; STREET: 345 Park Avenue

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy Disk

; COMPUTER: IBM PC Compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WORDPERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/07201

; FILING DATE: 06-JUN-1995

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/367,841

; FILING DATE: 30-DEC-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/257,963

; FILING DATE: 07-JUN-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/952,796

; FILING DATE: 24-SEP-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: DOROTHY R. AUTH

; REGISTRATION NUMBER: 36434

; REFERENCE/DOCKET NUMBER: 20264126PCT

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 758-4800

; TELEFAX: (212) 751-6849

; INFORMATION FOR SEQ ID NO: 10:

/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 7210 Base Pairs
/ TYPE: Nucleic Acid
/ STRANDEDNESS: Double
/ TOPOLOGY: Unknown
/ MOLECULE TYPE: Genomic DNA
/ ORIGINAL SOURCE:
/ ORGANISM: Human
/ IMMEDIATE SOURCE:
/ LIBRARY: DASH II
/ FEATURE:
/ NAME/KEY: JT6A
/ LOCATION:
/ IDENTIFICATION METHOD:
/ OTHER INFORMATION: 7.0 kb Not 1-Not
/ OTHER INFORMATION: fragment; Derived from human placental
/ OTHER INFORMATION: genomic DNA; also referred to as JT106
/ PCT-US95-07201-10

Query Match 27.8%; Score 277.8; DB 5; Length 7210;
Best Local Similarity 82.3%; Pred. No. 3.5e-54;
Matches 354; Conservative 0; Mismatches 73; Indels 3; Gaps 3;
QY 88 GCCTAGCGCGGTGGCTCAAGCCTGTAAATCCAGCAGCTTTAGAGGTCGAAGAGGTGGA 147
DB 3756 GCGCGGCGACGGTGGCTCAGCGCTGTAGTCCAGCAGCTTTGGGAGCCGAGCGAGGCGAGA 3815
QY 148 TCACCTGAGGTGAGGCTTTTGAGACCGCCTGGCGCAACCGGTGAACCCCAATCTCTAC 207
DB 3816 TCACCTGAGGTGAGGCTTTTGAGACCGCCTGGCGCAACCGGTGAACCCCAATCTCTAC 3874
QY 208 TAAATAATA-AAAAATAGCTNGGCTGGCTGCTACACCTGTAAATCCAGCAGCTTTGGG 266
DB 3875 TAAATAATA-AAAAATAGCTNGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3934
QY 267 AGCTGAGAGCGGTGGATCATCCTGAAGTCAGGAGTTCAGAGGCGAGCTGGGCAACATGTT 326
DB 3935 AGCAGAGGTGGCGAGATCCTTGAAGTCAGGAGTTCAGAGGCGAGCTGGGCAACATGTT 3994
QY 327 GAAACACGCTCTCTACTAAATAACAAATTAAGCAGGTCAGGAGTTCAGAGGCGAGCTGG 386
DB 3995 GAAACACGCTCTCTACTAAATAACAAATTAAGCAGGTCAGGAGTTCAGAGGCGAGCTGG 4054
QY 387 TCCAGCTACTTGGAGGCTGAGGCGGAGAGTTCAGGTCAGGAGTTCAGGAGGAGTTCG 446
DB 4055 TCCAGCTACTTGGAGGCTGAGGCGGAGAGTTCAGGTCAGGAGTTCAGGAGGAGTTCG 4114
QY 447 AGTGAGCGGAGATAGAGTCACTGACCTCCAGCTGAGGTCAGGAGGAGTTCGCTCTC 506
DB 4115 AGTGAGCGGAGATAGAGTCACTGACCTCCAGCTGAGGTCAGGAGGAGTTCGCTCTC 4173
QY 507 AGAAATAAAA 516
DB 4174 AAAAAAAAAA 4183

RESULT 9
US-08-520-373D-4
; Sequence 4, Application US/08520373D
; Patent No. 6451763
; GENERAL INFORMATION:
; APPLICANT: Tombran-Tink, Joyce
; APPLICANT: Steele, Fintan R
; APPLICANT: Chader, Gerald J
; APPLICANT: Becerra, Sofia P
; APPLICANT: Johnson, Lincoln V
; APPLICANT: Rodriguez, Ignacio R
; TITLE OF INVENTION: RETINAL PIGMENTED EPITHELIUM DERIVED NEUROTROPIC FACTOR
; FILE REFERENCE: 2026-4203US1
; CURRENT APPLICATION NUMBER: US/08/520,373D
; CURRENT FILING DATE: 1995-08-29
; PRIOR APPLICATION NUMBER: 08/377,710
; PRIOR FILING DATE: 1995-01-25

/ PRIOR APPLICATION NUMBER: 08/279,979
/ PRIOR FILING DATE: 1994-07-25
/ PRIOR APPLICATION NUMBER: 07/894,215
/ PRIOR FILING DATE: 1992-06-04
/ PRIOR APPLICATION NUMBER: 07/952,796
/ PRIOR FILING DATE: 1992-09-24
/ NUMBER OF SEQ ID NOS: 34
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 4
/ LENGTH: 14581
/ TYPE: DNA
/ ORGANISM: HUMAN
/ FEATURE:
/ OTHER INFORMATION: mRNA: 6683; EXON: 6683-6790; EXON 11584-11675;
/ OTHER INFORMATION: EXON: 14539-14581; INTRON: 6791-11583; INTRON:
/ OTHER INFORMATION: 11676-14538; CDS: 11584-11675; 14539-14580
/ US-08-520-373D-4

Query Match 27.8%; Score 277.8; DB 4; Length 14581;
Best Local Similarity 82.3%; Pred. No. 4e-54;
Matches 354; Conservative 0; Mismatches 73; Indels 3; Gaps 3;
QY 88 GCCTAGCGCGGTGGCTCAGCCTGTAAATCCAGCAGCTTTAGAGGTCGAAGAGGTGGA 147
DB 3755 GCGCGGCGACGGTGGCTCAGCGCTGTAGTCCAGCAGCTTTGGGAGCCGAGCGAGGCGAGA 3814
QY 148 TCACCTGAGGTGAGGCTTTTGAGACCGCCTGGCGCAACCGGTGAACCCCAATCTCTAC 207
DB 3815 TCACCTGAGGTGAGGCTTTTGAGACCGCCTGGCGCAACCGGTGAACCCCAATCTCTAC 3873
QY 208 TAAATAATA-AAAAATAGCTNGGCTGGCTGCTACACCTGTAAATCCAGCAGCTTTGGG 266
DB 3874 TAAATAATA-AAAAATAGCTNGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3933
QY 267 AGCTGAGAGCGGTGGATCATCCTGAAGTCAGGAGTTCAGAGGCGAGCTGGGCAACATGTT 326
DB 3934 AGCAGAGGTGGCGAGATCCTTGAAGTCAGGAGTTCAGAGGCGAGCTGGGCAACATGTT 3993
QY 327 GAAACACGCTCTCTACTAAATAACAAATTAAGCAGGTCAGGAGTTCAGAGGCGAGCTGG 386
DB 3994 GAAACACGCTCTCTACTAAATAACAAATTAAGCAGGTCAGGAGTTCAGAGGCGAGCTGG 4053
QY 387 TCCAGCTACTTGGAGGCTGAGGCGGAGAGTTCAGGTCAGGAGTTCAGGAGGAGTTCG 446
DB 4054 TCCAGCTACTTGGAGGCTGAGGCGGAGAGTTCAGGTCAGGAGTTCAGGAGGAGTTCG 4113
QY 447 AGTGAGCGGAGATAGAGTCACTGACCTCCAGCTGAGGTCAGGAGGAGTTCGCTCTC 506
DB 4114 AGTGAGCGGAGATAGAGTCACTGACCTCCAGCTGAGGTCAGGAGGAGTTCGCTCTC 4172
QY 507 AGAAATAAAA 516
DB 4173 AAAAAAAAAA 4182

RESULT 10
US-08-367-841A-43
; Sequence 43, Application US/08367841A
; Patent No. 6319687
; GENERAL INFORMATION:
; APPLICANT: Chader, Gerald J.; Rodriguez,
; APPLICANT: Ignacio R.; Mazuruk, Krzysztof;
; APPLICANT: Tombran-Tink, Joyce
; TITLE OF INVENTION: PIGMENT EPITHELIUM
; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan
; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA

```

1  ZIP: 10154
2
3  COMPUTER REGISTRATION FORM:
4
5  MEDIUM TYPE: Floppy Disk
6  COMPUTER: IBM PC Compatible
7  OPERATING SYSTEM: PC-DOS/MS-DOS
8  SOFTWARE: WORDPERFECT 5.1
9
10 CURRENT APPLICATION DATA:
11
12 APPLICATION NUMBER: US/08/367,841A
13 FILING DATE: 30-DEC-1994
14 CLASSIFICATION: 435
15
16 PRIOR APPLICATION DATA:
17
18 APPLICATION NUMBER: 08/257,963
19 FILING DATE: 07-JUN-1994
20
21 PRIOR APPLICATION DATA:
22
23 APPLICATION NUMBER: 07/952,796
24 FILING DATE: 24-SEP-1992
25
26 ATTORNEY/AGENT INFORMATION:
27
28 NAME: DOROTHY R. AUTH
29 REGISTRATION NUMBER: 36434
30 REFERENCE/DOCKET NUMBER: 20264126US2
31
32 TELECOMMUNICATION INFORMATION:
33
34 TELEPHONE: (212) 758-4800
35 TELEFAX: (212) 751-6849
36
37 INFORMATION FOR SEQ ID NO: 43:
38
39 SEQUENCE CHARACTERISTICS:
40
41 LENGTH: 22481 Base Pairs
42 TYPE: Nucleic Acid
43 STRANDEDNESS: Double
44
45 TOPOLOGY: Unknown
46
47 MOLECULE TYPE: Genomic DNA
48
49 FEATURE:
50
51 NAME/KEY: Pl-147
52
53 LOCATION:
54
55 IDENTIFICATION METHOD:
56
57 OTHER INFORMATION: full length genomic
58
59 OTHER INFORMATION: sequence for FEDE
60
61 US-08-367-841A-43

```

Query Match	27.8%	Score 277.8	DB 4	Length 22481
Best Local Similarity	82.3%	Pred. No. 4.3e-54		
Matches 354	Conservative 0	Mismatches 73	Indels 3	Gaps
QY 88	GCCTAGCGCGGTGGCTCAGCCGCTGTAA	PCCGACGACTTTAGAGAGTTCGAGAGGTGGA	147	
Db				
QY 3748	GGCCGGCAGCGTGGCTCAGCCGCTGTAG	TCCGACGACTTTGGGAGCGCGAGGCGAGCAGA	3807	
Db				
QY 148	TCACCTTCAGGTCAGGAGTTTGTAGACCG	AGCTTGGCCGCAACAGCGTGAACCCCTACTCTAC	207	
Db				
QY 3808	TCACCTTGAGTCCAGGAG-TTCGAGAC	CGACTGGCTTAACAGCTGAACCCCGCTCTCTAC	3866	
Db				
QY 208	TAAAAATA-AAAAATTAGCTNGGTCG	GTGGTCTACACCTCTAATCCGACACTTTGGG	266	
Db				
QY 3867	TAAAAATACAAAAATTAGCTGGG	CACGGTGGCTGTAAATCCGACACTTTGGG	3926	
Db				
QY 267	AGGCTGAGACGGGTGGATCACTTGA	AGTCAGGAGTTCAAGGCCAGCGCTGGGCCAACATG	326	
Db				
QY 3927	AGGCAGAGGTGGGCAGATCACTTT	GAGTTCAGAGTTTGAGACCGAGCTAGCCACATGGT	3986	
Db				
QY 327	GAACACCAAGCTCTCTACTACAA	ATAATTAGCAGGTGTGGTGGCACACGCTGTAG	386	
Db				
QY 3987	GAACCCCATCTCTACTAAAC	TACAAAAATAGCCGGGAGTGGTGGCAGCTGCTGTAA	4046	
Db				
QY 387	TCCCAGCTACTTTGGGAGCTTGAGG	CGGGAAGAAATCGCTTGAACCCAGTAGGAGAGTTGC	446	
Db				
QY 4047	TCCCGCCAGTTCAGGAGCTTGAGG	CAGGAAATCACTGGAATCTCGAGTGGAGTGGC	4106	
Db				
QY 447	AGTGAGCCGAGATAAGAGTCACT	TGCATCCAGCGCTGGGTGACAGACGACAGCTCCCTCTC	506	
Db				
QY 4107	AGTGAGCCGAGAT-GGTACCTCT	GTACTTCCAGCGCTGGGGGACAGAGTAGACTCCGCTCTC	4165	
Db				
QY 507	AGAAAAATAAA	516		
Db				
QY 4166	AAAAAATAAA	4175		
Db				

```

RESULT 11
PCT-US95-07201-43
; Sequence 43, Application PC/TUS9507201
; GENERAL INFORMATION:
; APPLICANT: Chader, Gerald J.; Becerra, Sofia
; APPLICANT: Patricia; Schwartz, Joan P.;
; APPLICANT: Taniwaki, Takayuki
; TITLE OF INVENTION: PIGMENT EPIETHILUM
; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
; TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PEDF GENE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07201
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/367,841
; FILING DATE: 30-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/257,963
; FILING DATE: 07-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/952,796
; FILING DATE: 24-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36434
; REFERENCE/DOCKET NUMBER: 20364126PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22481 Base Pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: Genomic DNA
; FEATURE:
; NAME/KEY: PL-147
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: full length genomic
; OTHER INFORMATION: sequence for PEDF plus flanking sequences.
PCT-US95-07201-43

```

	Query Match	27.8%;	Score 277.8;	DB 5;	Length 22481;
	Best Local Similarity	82.3%;	Pred. No. 4.3e-54;		
	Matches 354;	Conservative	0;	Mismatches 73;	Indels 3;
	Gaps	3;			
QY	88	GCCTAGCGCGTGGCTCAGCGCTCTAATCCAGCACTTTAAAGGTCGAAGAGGGTGA	147		
Db	3748	GCCTGGCAGCGTGGCTCAGCGCTCTAGTCCAGCACTTTGGAGCGCGAGCGCAGCA	3807		
QY	148	TCACTTGAGGTCAGGAGTTTGGAGCACCAGGCTGGCCACACGGGTGAACCCCATCTCTAC	207		
Db	3808	TCACTTGAGGTCAGGAG-TTCGAGACACAGCTGGCTTACACAGATGAACCCCGCTCTAC	3866		
QY	208	TAAAAATA-AAAAATTAGCTNGGTCGGGTGGCTCACACCTGTAAATCCAGCACATTTGGG	286		

Db 3867 TAAAAATACAAAAATAGCTGGCAGCGTGGCTGGTCCCTGTAATCCAGCAGCTTTGGG 3926
QY 267 AGGTGAGAGCGGTGGATCAGCTGAAGTCAGGAGTCAAGGCCAGCGCTGGGCAACATGCT 376
Db 3927 AGGCAGAGGTGGCGAGATCAGTTGAGGTCAGGAGTTGACAGCAGCGCTAGCCACATGGT 3986
QY 327 GAAACACAGCTCTACTAATAATACAAAAATAGCCAGGTTGGTGGCAGCAGCTGTAG 386
Db 3987 GAAACCCCAATCTCTACTAATAATACAAAAATAGCCAGGTTGGTGGCAGCTGTAG 4046
QY 387 TCCAGCTACTTTGGAGGCTGAGCGGAAGATCGCTGAACCCAGTAGGAGGAGGTGC 446
Db 4047 TCCAGCCAGCTCAGGAGGCTGAGCGAGAGATCAGTGGAAATCCTGGAGGTTGGAGTGC 4106
QY 447 AGTGAGCCAGATAGAGTCACTGCACCTCCAGCCTGGGTGACAGAGCAAGACTCCCTCTC 506
Db 4107 AGTGAGCCAGAT-GGTACCTCTGTACTCCAGCCTGGGAGAGAGTGGAGTCCGCTCTC 516
QY 507 AGAAAAATAAA 516
Db 4166 AAAAAAATAAA 4175

RESULT 12

US-09-875-223-2

; Sequence 2, Application US/09875223
; Patent No. 6391850
; GENERAL INFORMATION:
; APPLICANT: No. 6391850 Western University
; APPLICANT: David Bouck
; APPLICANT: Paul Gillis
; TITLE OF INVENTION: Methods and Compositions for Inhibiting Angiogenesis
; FILE REFERENCE: 0290-2303
; CURRENT APPLICATION NUMBER: US/09/875,223
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 09/122,079
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: PCT/US98/15228
; PRIOR FILING DATE: 1998-07-23
; PRIOR APPLICATION NUMBER: US 08/899,304
; PRIOR FILING DATE: 1997-07-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 22484
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: 1...22484
; OTHER INFORMATION: "n" means either a, c, t, or g
US-09-875-223-2

Query Match 27.8%; Score 277.8; DB 4; Length 22484;
Best Local Similarity 82.3%; Pred. No. 4.3e-54;
Matches 354; Conservative 0; Mismatches 73; Indels 3; Gaps 3;
QY 88 GGTAGGCGGCTGGCTCAGCGCTGTATCCAGCAGCTTTAGAGGTTCGAAGAGGTTGA 147
Db 3748 GCGCGGCGGCTGGCTCAGCGCTGTATCCAGCAGCTTTGGGAGCGCCAGGAGGAGA 3807
QY 148 TCACCTGAGTCAGGAGTTTGGAGCGCGCTGGCCACAGCGGTGAACCCATCTCTAC 207
Db 3808 TCACCTGAGTCAGGAG-TTCGAGACCGCGCTGGCTAACAGATGAACCCGCTCTAC 3866
QY 208 TAAAAATA-AAAAATAGCTNGGGTGGCTCAGACCTGTAAATCCAGCAGCTTTGGG 266
Db 3867 TAAAAATACAAAAATAGCTGGCAGCGGTGCTCGCTGTATATCCAGCAGCTTTGGG 3926
QY 267 AGGCTGAGAGCGGTGGATCAGCTGAAGTCAGGAGTTCAGGCCAGCCTGGGCAACATGGT 326

Db 3927 AGGCAGAGGTGGCGAGATCAGTTGAGTCAGGAGTTTGAGACCCAGCCTAGCCAAATGGT 3986
QY 327 GAAACACAGCTCTCTACTAATAATACAAAAATAGCCAGGTTGGTGGCAGCAGCTGTAG 386
Db 3987 GAAACCCCAATCTCTACTAATAATACAAAAATAGCCAGGTTGGTGGCAGCAGCTGTAG 4046
QY 387 TCCAGCTACTTTGGAGGCTGAGCGGAAGATCGCTGAACCCAGTAGGAGGAGGTGC 446
Db 4047 TCCAGCCAGCTCAGGAGGCTGAGCGAGAGATCAGTGGAAATCCTGGAGTGGAGTGC 4106
QY 447 AGTGAGCCAGATAGAGTCACTGCACCTCCAGCCTGGGTGACAGAGCAAGACTCCCTCTC 506
Db 4107 AGTGAGCCAGAT-GGTACCTCTGTACTCCAGCCTGGGAGAGAGTGGAGTCCGCTCTC 516
QY 507 AGAAAAATAAA 516
Db 4166 AAAAAAATAAA 4175

RESULT 13

US-09-741-150-3
; Sequence 3, Application US/09741150
; Patent No. 6436689
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000968
; CURRENT APPLICATION NUMBER: US/09/741,150
; CURRENT FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 112132
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(112132)
; OTHER INFORMATION: n = A,T,C or G
US-09-741-150-3

Query Match 27.6%; Score 275.6; DB 4; Length 112132;
Best Local Similarity 79.9%; Pred. No. 1.8e-53;
Matches 373; Conservative 0; Mismatches 90; Indels 4; Gaps 4;
QY 89 GCTAGGCGGCTGGCTCAGCGCTGTATCCAGCAGCTTTAGAGGTTCGAAGAGGTTGAT 148
Db 90602 GCTGGAGCTGGTGGCTCAGCAGCTGTAAATCCCAACACTATATGGGAAGCTGAGCGAGCAGAT 90661
QY 149 CACTTGAGTCAAGAGTTTGGAGACCGCTGGCCACACAGCTGAACCCCATCTCTACT 208
Db 90662 CACTTGAGTCAAGAG-TTCAGACCCAGCTGGTCAACATGTGTAACCCCGCTCTACT 90720
QY 209 AAAAAATA-AAAAATAGCTNGGGTGGCTCAGACCTGTAAATCCAGCAGCTTTGGGA 267
Db 90721 AAAAAATACAAATTTAGCCAGGCTCAGTGGCTCAGACCTGTAAATCCAGCAGTTTGGGA 90780
QY 268 GGCCTGAGCGGTGGATCAGCTGAAGTCAGGAGTTCAGGCCAGCCTGGGCAACATGGTG 327
Db 90781 GGCCTGAGTGGCGGATCAGCTGAGTTCAGGAGTTCAGGCCAGCCTGGGCAACATGGTG 90840
QY 328 AAACACAGCTCTCTACTAATAATAACAAAAATAGCCAGGTTGAGGAGTTCAGGCCAGCTGTAGT 387
Db 90841 AAACCCCAATCTCTACTAATAATAACAAAAATAGCTGGGATGTGGGATGCTGTAGT 90900
QY 388 CCCAGCTACTTTGGAGGCTGAGCGGAAGATCGCTTGAACCCAGTAGGAGGAGGTTGCA 447
Db 90901 CCCAGCTAC-AGGAGGTTGAGCCAGGAGATCGCTTGAACCCAGGAGGTTGGAGGTTGCA 90959
QY 448 GTGAGCCAGATAGAGTCACTGCACCTCCAGCCTGGGTGACAGAGCAAGACTCCCTCTCA 507


```

; FILE REFERENCE: CL001158
; CURRENT APPLICATION NUMBER: US/09/800,960
; CURRENT FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 62804
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc. feature
; LOCATION: (1..62804)
; OTHER INFORMATION: n = A,T,C or G
US-09-800-960-3

Query Match          27.1%; Score 271.4; DB 4; Length 62804;
Best Local Similarity 80.1%; Pred. No. 1.5e-52;
Matches 367; Conservative 0; Mismatches 87; Indels 4; Gaps

QY 79 AATGGATAGGCTAGGCGCGGTGGCTCAGCGCTGTAAATCCGACGACTTTAGAGGTGCAA 138
Db 53984 AAAATAATAGCCAGGTGCGGTGACTCACACGTGTAATCCGACACTTTGGGAGCGGAG 54043
QY 139 GAGGGTGGATCACTTTGAGGTCAGAGTATTTTGAGACAGCGCTGGCCACACGCTGAACCC 198
Db 54044 GCAAGTGGATCACTTGAGGTCAGAG-TTCAAAACCGCGCTGGCCACACATGATGAACCC 54102
QY 139 CATCTCTACTAAATAATAAAATTTAGTGTGGTGGCTGCGGTGCCTCACACCTGTAAATCCCGAGC 258
Db 54103 CGTCTCTACTAAATAATAAAAAAGCGCGGGGTAGTGGCTCACGCTGTAAATCCCGAAC 54162
QY 259 ACTTTGGGAGGTCAGACGGGTGGATCACTTGAA-GTCAAGAGTTTCAAGGCCAGCGCTGGG 317
Db 54163 ACTTTGGGAGCGCAAGGTGGGTGGATCACTTGAAGGTCAAGGTCAGAAGTTCAAGCACCGCTGGC 54222
QY 318 CAACATGGTGAACACCGTCTCTACTAAAAATACAAAAATTTAGCCAGGTGTGGTGGCACA 377
Db 54223 CAACATGGTGAACCTCCATCTCTACTAAAAATATAAAAAATTAGCCAGGTGTGGGCGAGG 54282
QY 378 CGCCTCTAGTCCGAGCTACTTTGGAGGCTCAGGCGGGAAGATCGCTTGAACCCAGTAGGC 437
Db 54283 TGCCTCTAATCGTAGTACTTCGGGAGGCGGAGGTGGGGAATCGCTTGAACCTGGGAGGT 54342
QY 438 AGAGGTTGCAGTGAGCGGAGATAGAGTGCTACTGCACTCCAGCCTGGGTGAC-AGAGCAAG 496
Db 54343 GGAGGTTGCAGTGAGCGGAGATCA-CCCCATTGCACTTCCAGCCTGGGCACAGAGCGAA 54401
QY 497 ACTCCCTCTCAGAAAATAAATAATAATAAATAAATA 534
Db 54402 ACTTCTCTCAAAAAAATAAATAAATAAATAAATAAATA 54439

```

Search completed: November 24, 2002, 09:22:16
Job time : 322 secs

```

RESULT 15
US-09-800-960-3
; Sequence 3, Application US/09800960
; Patent No. 6387677
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2002, 12:00:58 ; Search time 99 Seconds
(without alignments)
3825.558 Million cell updates/sec

Title: US-09-784-423-32
Perfect score: 1000
Sequence: 1 GGTGTGACCTTATCTCTCT.....CAGATGATGACCGCGTGC 1000

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 335578 seqs, 189365133 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671156

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1000	100.0	1000	10	US-09-784-423-32
2	52	5.2	2701	10	US-09-784-877-3378
3	52	5.2	1503841	9	US-09-946-807-1
4	52	5.2	1503841	10	US-09-795-668-1
5	52	5.2	1503841	10	US-09-795-686-1
6	50	5.0	276	10	US-09-764-869-2086
7	50	5.0	545	10	US-09-741-148A-30
8	50	5.0	10503	10	US-09-764-870-579
9	50	5.0	10503	10	US-09-764-853-854
10	50	5.0	10503	10	US-09-764-860-1032
11	50	5.0	13819	10	US-09-764-877-2596
12	50	5.0	23668	10	US-09-741-148A-3
13	50	5.0	26928	10	US-09-880-107-2278
14	50	5.0	111282	12	US-10-09A-989-3
15	50	5.0	143306	10	US-09-729-920-3
16	49	4.9	401	9	US-09-946-807-806
17	49	4.9	401	10	US-09-795-668-806
18	49	4.9	401	10	US-09-795-686-806
19	49	4.9	24533	9	US-09-764-868-1349
					Sequence 32, Appl
					Sequence 3378, Ap
					Sequence 1, Appli
					Sequence 1, Appli
					Sequence 2086, Ap
					Sequence 30, Appl
					Sequence 579, App
					Sequence 854, App
					Sequence 1032, Ap
					Sequence 2596, Ap
					Sequence 3, Appli
					Sequence 2278, Ap
					Sequence 3, Appli
					Sequence 806, App
					Sequence 806, App
					Sequence 1349, Ap

20	49	4.9	99014	10	US-09-880-107-3428	Sequence 3428, Ap
21	49	4.9	99014	10	US-09-880-107-3428	Sequence 3428, Ap
22	48	4.8	1514	10	US-09-822-849A-510	Sequence 510, App
23	48	4.8	180216	10	US-09-835-232-6	Sequence 6, Appli
24	47	4.7	337	10	US-09-764-869-1528	Sequence 1528, Ap
25	47	4.7	418	10	US-09-867-701-8320	Sequence 8320, Ap
26	47	4.7	546	10	US-09-998-598-647	Sequence 647, App
27	47	4.7	562	10	US-09-764-877-2628	Sequence 2628, Ap
28	47	4.7	4766	9	US-09-764-868-1443	Sequence 1443, Ap
29	47	4.7	6124	8	US-08-913-322-21	Sequence 21, Appl
30	47	4.7	6124	10	US-09-967-768A-184	Sequence 184, App
31	47	4.7	6133	8	US-08-913-322-2	Sequence 2, Appli
32	47	4.7	6228	8	US-08-913-322-23	Sequence 23, Appl
33	47	4.7	10867	10	US-09-764-869-2113	Sequence 2113, Ap
34	47	4.7	10867	10	US-09-764-877-3970	Sequence 3970, Ap
35	47	4.7	10894	10	US-09-764-869-2111	Sequence 2111, Ap
36	47	4.7	10907	10	US-09-764-869-2112	Sequence 2112, Ap
37	47	4.7	10907	10	US-09-764-877-3968	Sequence 3968, Ap
38	47	4.7	12718	10	US-09-764-877-3972	Sequence 3972, Ap
39	47	4.7	56737	10	US-09-782-378A-17	Sequence 17, Appl
40	47	4.7	68804	10	US-09-740-041-3	Sequence 3, Appli
41	47	4.7	84539	10	US-09-962-436-36	Sequence 36, Appl
42	46	4.6	574	10	US-09-764-847-1713	Sequence 1713, Ap
43	46	4.6	574	10	US-09-764-847-1714	Sequence 1714, Ap
44	46	4.6	2939	9	US-09-764-868-1314	Sequence 1314, Ap
45	46	4.6	2939	9	US-09-764-868-1315	Sequence 1315, Ap

ALIGNMENTS

RESULT 1

US-09-784-423-32

Sequence 32, Application US/09784423

Patent No. US20020012924A1

GENERAL INFORMATION:

APPLICANT: Schumm, James W.

Bacher, Jeffery W.

TITLE OF INVENTION: MATERIALS AND METHODS FOR

IDENTIFYING AND ANALYZING

REPEAT DNA MARKERS

NUMBER OF SEQUENCES: 147

CORRESPONDENCE ADDRESS:

ADDRESSEE: Promega Corporation

STREET: 2800 Woods Hollow Road

City: Madison

STATE: Wisconsin

COUNTRY: U.S.A.

ZIP: 53711-5399

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette - 3.5 inch, 1.44 MB

COMPUTER: IBM compatible PC

OPERATING SYSTEM: Windows 95

SOFTWARE: Word 97 (DOS text format)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/784,423

FILING DATE: 15-Feb-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/018,584

FILING DATE: 04-Feb-1998

ATTORNEY/AGENT INFORMATION:

NAME: Grady J. Frenchick

REGISTRATION NUMBER: 29,018

REFERENCE/DOCKET NUMBER: 16026.9180

TELECOMMUNICATION INFORMATION:

TELEPHONE: (608) 257-3501

TELEFAX: (608) 257-2275

INFORMATION FOR SEQ ID NO: 32

SEQUENCE CHARACTERISTICS:

LENGTH: 1000 bp

TYPE: Nucleic Acid

STRANDEDNESS: Double


```

; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: k-g or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: s-g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: w-a or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: b-g or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: d-a or g or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: h-a or c or t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: v-a or g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: n-a or g or c or t/u
;
US-09-946-807-1

```

```

Query Match          5.2%; Score 52; DB 9; Length 1503841;
Best Local Similarity 100.0%; Pred. No. 3.5e-15;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 509 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATTC 1366616
|||||
Db 1366667 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATTC 1366616

```

```

RESULT 4
US-09-795-668-1/c
; Sequence 1, Application US/09795668
; Patent No. US20020045577A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn
; APPLICANT: Steinthorsdottir, Valgerdur
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; CURRENT APPLICATION NUMBER: US/09/795,668
; PRIOR FILING DATE: 2001-02-28
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1503841
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: r-g or a
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: y-t/u or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: m-a or c
; NAME/KEY: misc_feature

```

```

; LOCATION: (1)...(1531)
; OTHER INFORMATION: k-g or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: s-g or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: w-a or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: b-g or c or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: d-a or g or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: h-a or c or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: v-a or g or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: n-a or g or c or t/u
;
US-09-795-668-1

Query Match          5.2%; Score 52; DB 10; Length 1503841;
Best Local Similarity 100.0%; Pred. No. 3.5e-15;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 509 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATTC 1366616
|||||
Db 1366667 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATTC 1366616

RESULT 5
US-09-795-686-1/c
; Sequence 1, Application US/09795686
; Patent No. US20020094954A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn
; APPLICANT: Steinthorsdottir, Valgerdur
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345.2005-001
; CURRENT APPLICATION NUMBER: US/09/795,686
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/515,715
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1503841
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: r-g or a
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: y-t/u or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: m-a or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: k-g or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: s-g or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: w-a or t/u

```

1

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1032
; LENGTH: 10503
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-860-1032

Query Match 5.0%; Score 50; DB 10; Length 10503;
Best Local Similarity 100.0%; Pred. No. 3.1e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 166 TTGTGACGACGCTGGCCACACGGGTGAACCCCATCTCTACTAAAAATA 215
|||||
DB 4291 TTGTGACGACGCTGGCCACACGGGTGAACCCCATCTCTACTAAAAATA 4340

RESULT 11
US-09-764-877-2596/c
; Sequence 2596, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2596
; LENGTH: 13819
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2596

Query Match 5.0%; Score 50; DB 10; Length 13819;
Best Local Similarity 100.0%; Pred. No. 3.1e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 509 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 558
|||||
DB 3482 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 3433

RESULT 12
US-09-741-148A-3/c
; Sequence 3, Application US/09741148A
; Patent No. US20020076750A1
; GENERAL INFORMATION:
; APPLICANT: Chunhua YAN et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000566
; CURRENT APPLICATION NUMBER: US/09/741,148A
; CURRENT FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: 60/206,982
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 23668
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature

; LOCATION: (1)...(23668)
; OTHER INFORMATION: n = A,T,C or G
US-09-741-148A-3

Query Match 5.0%; Score 50; DB 10; Length 23668;
Best Local Similarity 100.0%; Pred. No. 3.1e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 509 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 558
|||||
DB 9671 AAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 5622

RESULT 13
US-09-880-107-2278
; Sequence 2278, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherif, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2278
; LENGTH: 26928
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M17262
US-09-880-107-2278

Query Match 5.0%; Score 50; DB 10; Length 26928;
Best Local Similarity 100.0%; Pred. No. 3.1e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 336 TCCTCTACTAAATACAAAAATAGCCAGGTGTGTGGCACACGCTGTA 385
|||||
DB 5015 TCCTCTACTAAATACAAAAATAGCCAGGTGTGTGGCACACGCTGTA 5064

RESULT 14
US-10-094-989-3
; Sequence 3, Application US/10094989
; Patent No. US20020115179A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHODIESTERASE
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: PHOSPHODIESTERASE PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001063DIV
; CURRENT APPLICATION NUMBER: US/10/094,989
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/754,250
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 111282
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(111282)
; OTHER INFORMATION: n = A,T,C or G

; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1349
; LENGTH: 24533
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-868-1349

Query Match 4.9%; Score 49; DB 9; Length 24533;
Best Local Similarity 100.0%; Pred. No. 9.1e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 510 AAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 558
|||||
Db 5926 AAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 5878

RESULT 20

US-09-880-107-3428
; Sequence 3428, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3428
; LENGTH: 99014
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U78027
US-09-880-107-3428

Query Match 4.9%; Score 49; DB 10; Length 99014;
Best Local Similarity 100.0%; Pred. No. 9.1e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 337 CTTCTACTAAATAACAAAATTAGCCAGGTGTGGTGGCAGCGCTGTA 385
|||||
Db 43211 CTTCTACTAAATAACAAAATTAGCCAGGTGTGGTGGCAGCGCTGTA 43259

RESULT 21

US-09-880-107-3428/c
; Sequence 3428, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3428

; LENGTH: 99014
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U78027
US-09-880-107-3428

Query Match 4.9%; Score 49; DB 10; Length 99014;
Best Local Similarity 100.0%; Pred. No. 9.1e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 510 AAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 558
|||||
Db 98924 AAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 98876

RESULT 22

US-09-822-849A-510/c
; Sequence 510, Application US/09822849A
; Patent No. US20020045170A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fachtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakara
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6403
; CURRENT APPLICATION NUMBER: US/09/822,849A
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/195,582
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 598
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 510
; LENGTH: 1514
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-849A-510

Query Match 4.8%; Score 48; DB 10; Length 1514;
Best Local Similarity 100.0%; Pred. No. 2.7e-13;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 510 AAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 557
|||||
Db 1451 AAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAAT 1404

RESULT 23

US-09-835-232-6/c
; Sequence 6, Application US/09835232
; Patent No. US20020098489A1
; GENERAL INFORMATION:
; APPLICANT: Leder, Philip
; APPLICANT: Leader, Benjamin
; TITLE OF INVENTION: FORMIN-2 NUCLEIC ACIDS AND POLYPEPTIDES
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 00383/052002
; CURRENT APPLICATION NUMBER: US/09/835,232
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: US 60/196,811
; PRIOR FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 180216
; TYPE: DNA
; ORGANISM: Homo sapiens


```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(180216)
; OTHER INFORMATION: n = A,T,C or G
US-09-835-232-6

Query Match
; Sequence 48; Score 48; DB 10; Length 180216;
; Best Local Similarity 100.0%; Pred. No. 2.7e-13;
; Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 369 GGTGGCACACCCCTGTAGTCCAGCTACTTGGGAGGCTGAGCGCGAAG 416
|||||
Db 9148 GTGGCACACCCCTGTAGTCCAGCTACTTGGGAGGCTGAGCGCGAAG 9101

RESULT 24
US-09-764-869-1528
; Sequence 1528, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1528
; LENGTH: 337
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-1528

Query Match
; Sequence 47; Score 47; DB 10; Length 337;
; Best Local Similarity 100.0%; Pred. No. 8.1e-13;
; Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 245 CTTGTAATCCACACTTTGGGAGGCTGAGACGGGTGGATCACCTGA 291
|||||
Db 76 CCTGTAATCCACACTTTGGGAGGCTGAGACGGGTGGATCACCTGA 122

RESULT 25
US-09-867-701-8320
; Sequence 8320, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8320
; LENGTH: 418
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(418)
; OTHER INFORMATION: n = A,T,C or G
US-09-867-701-8320

Query Match
; Sequence 47; Score 47; DB 10; Length 418;
; Best Local Similarity 100.0%; Pred. No. 8.1e-13;
; Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GGGTGGGTGGCTCACACCTGTATATCCAGCACCTTTGGGAGGCTGAG 274
```

```
|||||
Db 143 GGGTGGGTGGCTCACACCTGTATATCCAGCACCTTTGGGAGGCTGAG 189

RESULT 26
US-09-998-598-647/c
; Sequence 647, Application US/09998598
; Patent No. US20020150922A1
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Chenault, Ruth A.
; APPLICANT: Meagher, Madelein Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.561
; CURRENT APPLICATION NUMBER: US/09/998,598
; CURRENT FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 2606
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 647
; LENGTH: 546
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-998-598-647

Query Match
; Sequence 47; Score 47; DB 10; Length 546;
; Best Local Similarity 100.0%; Pred. No. 8.1e-13;
; Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GGGTGGGTGGCTCACACCTGTATATCCAGCACCTTTGGGAGGCTGAG 274
|||||
Db 89 GGGTGGGTGGCTCACACCTGTATATCCAGCACCTTTGGGAGGCTGAG 43

RESULT 27
US-09-764-877-2628
; Sequence 2628, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2628
; LENGTH: 562
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2628

Query Match
; Sequence 47; Score 47; DB 10; Length 562;
; Best Local Similarity 100.0%; Pred. No. 8.1e-13;
; Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GGGTGGGTGGCTCACACCTGTATATCCAGCACCTTTGGGAGGCTGAG 274
|||||
Db 207 GGGTGGGTGGCTCACACCTGTATATCCAGCACCTTTGGGAGGCTGAG 253

RESULT 28
US-09-764-868-1443
; Sequence 1443, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
```

```
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1443
; LENGTH: 4766
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-868-1443

Query Match          4.7%; Score 47; DB 9; Length 4766;
Best Local Similarity 100.0%; Pred. No. 8e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GGGTGGGTGGCTCACACCTGTAATCCAGCAGCTTTGGGAGGCTGAG 274
      |||||||
Db 2594 GGGTGGGTGGCTCACACCTGTAATCCAGCAGCTTTGGGAGGCTGAG 2640

RESULT 29
US-08-913-322-21
; Sequence 21, Application US/08913322
; Patent No. US20020137028A1
; GENERAL INFORMATION:
; APPLICANT: Korneluk, Robert G.
; APPLICANT: MacKenzie, Alexander E.
; APPLICANT: Roy, Natalie
; APPLICANT: Robertson, George
; APPLICANT: Tamai, Katsu
; TITLE OF INVENTION: USER OF NEURONAL APOPTOSIS INHIBITOR
; FILE REFERENCE: 07891/013001
; CURRENT APPLICATION NUMBER: US/08/913,322
; EARLIER FILING DATE: 1997-09-12
; EARLIER FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: PCT/IB97/00142
; EARLIER FILING DATE: 1996-01-19
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 6124
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-913-322-21

Query Match          4.7%; Score 47; DB 8; Length 6124;
Best Local Similarity 100.0%; Pred. No. 8e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GGGTGGGTGGCTCACACCTGTAATCCAGCAGCTTTGGGAGGCTGAG 274
      |||||||
Db 5487 GGGTGGGTGGCTCACACCTGTAATCCAGCAGCTTTGGGAGGCTGAG 5533

RESULT 30
US-09-967-768A-184
; Sequence 184, Application US/09967768A
; Patent No. US20020150877A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-72
; CURRENT APPLICATION NUMBER: US/09/967,768A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/60/236,109
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,034
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,111
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: PatentIn version 3.0
```

```
; SEQ ID NO 184
; LENGTH: 6124
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-967-768A-184

Query Match          4.7%; Score 47; DB 10; Length 6124;
Best Local Similarity 100.0%; Pred. No. 8e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GGGTGGGTGGCTCACACCTGTAATCCAGCAGCTTTGGGAGGCTGAG 274
      |||||||
Db 5487 GGGTGGGTGGCTCACACCTGTAATCCAGCAGCTTTGGGAGGCTGAG 5533

RESULT 31
US-08-913-322-2
; Sequence 2, Application US/08913322
; Patent No. US20020137028A1
; GENERAL INFORMATION:
; APPLICANT: Korneluk, Robert G.
; APPLICANT: MacKenzie, Alexander E.
; APPLICANT: Roy, Natalie
; APPLICANT: Robertson, George
; APPLICANT: Tamai, Katsu
; TITLE OF INVENTION: USER OF NEURONAL APOPTOSIS INHIBITOR
; FILE REFERENCE: 07891/013001
; CURRENT APPLICATION NUMBER: US/08/913,322
; CURRENT FILING DATE: 1997-09-12
; EARLIER FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: PCT/IB97/00142
; EARLIER FILING DATE: 1996-01-19
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 6133
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-913-322-2

Query Match          4.7%; Score 47; DB 8; Length 6133;
Best Local Similarity 100.0%; Pred. No. 8e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GGGTGGGTGGCTCACACCTGTAATCCAGCAGCTTTGGGAGGCTGAG 274
      |||||||
Db 5496 GGGTGGGTGGCTCACACCTGTAATCCAGCAGCTTTGGGAGGCTGAG 5542

RESULT 32
US-08-913-322-23
; Sequence 23, Application US/08913322
; Patent No. US20020137028A1
; GENERAL INFORMATION:
; APPLICANT: Korneluk, Robert G.
; APPLICANT: MacKenzie, Alexander E.
; APPLICANT: Roy, Natalie
; APPLICANT: Robertson, George
; APPLICANT: Tamai, Katsu
; TITLE OF INVENTION: USER OF NEURONAL APOPTOSIS INHIBITOR
; FILE REFERENCE: 07891/013001
; CURRENT APPLICATION NUMBER: US/08/913,322
; CURRENT FILING DATE: 1997-09-12
; EARLIER FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: PCT/IB97/00142
; EARLIER FILING DATE: 1996-01-19
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
```

```
; LENGTH: 6228
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-913-322-23

Query Match
4.7%; Score 47; DB 8; Length 6228;
Best Local Similarity 100.0%; Pred. No. 8e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GGGTGGGTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAG 274
|||||
Db 5591 GGGTGGGTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAG 5637

RESULT 33
US-09-764-869-2113
; Sequence 2113, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2113
; LENGTH: 10867
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-2113

Query Match
4.7%; Score 47; DB 10; Length 10867;
Best Local Similarity 100.0%; Pred. No. 8e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GGGTGGGTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAG 274
|||||
Db 1100 GGGTGGGTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAG 1146

RESULT 34
US-09-764-877-3970
; Sequence 3970, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3970
; LENGTH: 10867
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3970

Query Match
4.7%; Score 47; DB 10; Length 10867;
Best Local Similarity 100.0%; Pred. No. 8e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GGGTGGGTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAG 274
|||||
Db 1100 GGGTGGGTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAG 1146

RESULT 35
US-09-764-869-2111
; Sequence 2111, Application US/09764869
```

```
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2111
; LENGTH: 10894
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-2111

Query Match
4.7%; Score 47; DB 10; Length 10894;
Best Local Similarity 100.0%; Pred. No. 8e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GGGTGGGTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAG 274
|||||
Db 1138 GGGTGGGTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAG 1184

RESULT 36
US-09-764-869-2112
; Sequence 2112, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2112
; LENGTH: 10907
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-2112

Query Match
4.7%; Score 47; DB 10; Length 10907;
Best Local Similarity 100.0%; Pred. No. 8e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GGGTGGGTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAG 274
|||||
Db 1137 GGGTGGGTGGCTCACACCTGTATCCAGCAGCTTTGGGAGGCTGAG 1183

RESULT 37
US-09-764-877-3968
; Sequence 3968, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3968
; LENGTH: 10907
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3968

Query Match
4.7%; Score 47; DB 10; Length 10907;
```

Best Local Similarity 100.0%; Pred. No. 8e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GGGTGGGTGGCTCACACCTGTAAATCCAGCAGCTTTGGGAGGCTGAG 274
|||||
Db 1137 GGGTGGGTGGCTCACACCTGTAAATCCAGCAGCTTTGGGAGGCTGAG 1183
|||||

RESULT 38

US-09-764-877-3972
; Sequence 3972, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3972
; LENGTH: 12718
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3972

Query Match 4.7%; Score 47; DB 10; Length 12718;
Best Local Similarity 100.0%; Pred. No. 8e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GGGTGGGTGGCTCACACCTGTAAATCCAGCAGCTTTGGGAGGCTGAG 274
|||||
Db 1137 GGGTGGGTGGCTCACACCTGTAAATCCAGCAGCTTTGGGAGGCTGAG 1183
|||||

RESULT 39

US-09-782-378A-17/c
; Sequence 17, Application US/09782378A
; Patent No. US20020102731A1
; GENERAL INFORMATION:
; APPLICANT: Hearing, Patrick
; APPLICANT: Bahou, Wadie
; APPLICANT: Sandalon, Ziv
; APPLICANT: Gnatenko, Dmitri
; TITLE OF INVENTION: Adenoviral Vectors
; FILE REFERENCE: STONYB-04970
; CURRENT APPLICATION NUMBER: US/09/782,378A
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/237,747
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 56737
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-378A-17

Query Match 4.7%; Score 47; DB 10; Length 56737;
Best Local Similarity 100.0%; Pred. No. 7.9e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 510 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 556
|||||
Db 18486 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 18440
|||||

RESULT 40

US-09-740-041-3/c
; Sequence 3, Application US/09740041
; Patent No. US20020082190A1
; GENERAL INFORMATION:

; APPLICANT: MERKULOV, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001001
; CURRENT APPLICATION NUMBER: US/09/740,041
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 66804
; TYPE: DNA
; ORGANISM: Human
US-09-740-041-3

Query Match 4.7%; Score 47; DB 10; Length 66804;
Best Local Similarity 100.0%; Pred. No. 7.9e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 GGGTGGGTGGCTCACACCTGTAAATCCAGCAGCTTTGGGAGGCTGAG 274
|||||
Db 35282 GGGTGGGTGGCTCACACCTGTAAATCCAGCAGCTTTGGGAGGCTGAG 35236
|||||

RESULT 41

US-09-962-436-36/c
; Sequence 36, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 36
; LENGTH: 84539
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-36

Query Match 4.7%; Score 47; DB 10; Length 84539;
Best Local Similarity 100.0%; Pred. No. 7.9e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 GAGACAGCCTGGCCACACAGCGGTGAACCCCATCTCTACTATAAATA 215
|||||
Db 10850 GAGACAGCCTGGCCACACAGCGGTGAACCCCATCTCTACTATAAATA 10804
|||||

RESULT 42

US-09-764-847-1713/c
; Sequence 1713, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1713
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:
; NAME/KEY: SITE
; LOCATION: (13)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-847-1713

Query Match 4.6%; Score 46; DB 10; Length 574;
Best Local Similarity 100.0%; Pred. No. 2.4e-12;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 170 AGACACCTGGCCACACGGTGAACCCCATCTCTACTAAATA 215
|||||
Db 219 AGACACCTGGCCACACGGTGAACCCCATCTCTACTAAATA 174

RESULT 43
US-09-764-847-1714/c
; Sequence 1714, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1714
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1714

Query Match 4.6%; Score 46; DB 10; Length 574;
Best Local Similarity 100.0%; Pred. No. 2.4e-12;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 170 AGACACCTGGCCACACGGTGAACCCCATCTCTACTAAATA 215
|||||
Db 219 AGACACCTGGCCACACGGTGAACCCCATCTCTACTAAATA 174

RESULT 44
US-09-764-868-1314
; Sequence 1314, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1314
; LENGTH: 2939
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-868-1314

Query Match 4.6%; Score 46; DB 9; Length 2939;
Best Local Similarity 100.0%; Pred. No. 2.4e-12;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 229 GTGGGGTGGCTCACACCTGTAAATCCAGCACTTTGGAGGCTGAG 274
|||||
Db 1942 GTGGGGTGGCTCACACCTGTAAATCCAGCACTTTGGAGGCTGAG 1987

RESULT 45
US-09-764-868-1315/c

; Sequence 1315, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1315
; LENGTH: 2939
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-868-1315

Query Match 4.6%; Score 46; DB 9; Length 2939;
Best Local Similarity 100.0%; Pred. No. 2.4e-12;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 229 GTGGGGTGGCTCACACCTGTAAATCCAGCACTTTGGAGGCTGAG 274
|||||
Db 998 GTGGGGTGGCTCACACCTGTAAATCCAGCACTTTGGAGGCTGAG 953

Search completed: November 24, 2002, 13:38:45
Job time : 1912 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
OM nucleic - nucleic search, using sw model
Run on: November 24, 2002, 13:30:43 ; Search time 46.5 seconds
(without alignments)
164.880 Million cell updates/sec
Title: US-09-784-423-124
Perfect score: 25
Sequence: 1 GGTTCAGTGAGCGGAGATAGAGT 25
Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0
Searched: 441362 seqs, 153338381 residues
Word size : 0
Total number of hits satisfying chosen parameters: 882724
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Listing first 45 summaries
Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	100.0	25	4	US-09-018-584A-124
2	25	100.0	1000	4	US-09-018-584A-32
3	20	80.0	764	4	US-09-288-143-57
4	20	80.0	1875	2	US-08-683-743-3
5	20	80.0	4042	4	US-08-406-030A-17
6	20	80.0	4129	2	US-08-370-319C-12
7	20	80.0	4129	4	US-09-224-834-12
8	20	80.0	45716	4	US-08-965-048-5
9	20	80.0	45989	4	US-08-965-048-6
10	19	76.0	21	1	US-08-133-629-2
c 11	19	76.0	239	2	US-08-687-080-93
c 12	19	76.0	265	2	US-08-849-701-1
c 13	19	76.0	294	2	US-08-481-658B-61
c 14	19	76.0	294	2	US-08-477-504A-61
c 15	19	76.0	294	2	US-08-486-756A-61
c 16	19	76.0	294	2	US-08-485-862B-61
c 17	19	76.0	294	3	US-08-787-739-61
c 18	19	76.0	294	3	US-08-487-077A-61
c 19	19	76.0	294	3	US-08-485-863A-61
c 20	19	76.0	294	4	US-08-485-049D-61
c 21	19	76.0	294	4	US-09-178-115-61
c 22	19	76.0	294	4	US-09-177-776-61
c 23	19	76.0	302	2	US-08-849-701-3
c 24	19	76.0	336	4	US-09-385-982-17
c 25	19	76.0	374	4	US-09-385-982-135
c 26	19	76.0	379	3	US-09-157-177-134
c 27	19	76.0	577	4	US-09-227-357-92

c 28	19	76.0	609	4	US-09-385-982-291
c 29	19	76.0	618	4	US-09-385-982-218
c 30	19	76.0	649	4	US-09-535-008-49
c 31	19	76.0	689	4	US-09-105-542A-14
c 32	19	76.0	696	4	US-09-740-235-16
c 33	19	76.0	775	4	US-09-227-357-108
c 34	19	76.0	789	2	US-08-394-152A-42
c 35	19	76.0	821	4	US-09-342-681C-7
c 36	19	76.0	825	2	US-08-486-148B-1
c 37	19	76.0	841	5	PCT-US93-06251-80
c 38	19	76.0	841	5	PCT-US93-06251-81
c 39	19	76.0	866	4	US-09-257-179-11
c 40	19	76.0	885	4	US-09-288-143-55
c 41	19	76.0	889	1	US-08-832-883-52
c 42	19	76.0	889	2	US-08-832-877-52
c 43	19	76.0	891	4	US-09-247-155-141
c 44	19	76.0	1000	4	US-09-018-584A-40
c 45	19	76.0	1000	4	US-09-641-638-636

ALIGNMENTS

RESULT 1
US-09-018-584A-124
; Sequence 124, Application US/09018584A
; Patent No. 6238863
; GENERAL INFORMATION:
; APPLICANT: Schumm, James W.
; APPLICANT: Bacher, Jeffery W.
; TITLE OF INVENTION: MATERIALS AND METHODS FOR IDENTIFYING AND ANALYZING INTERMEDIATE TANDEM
; TITLE OF INVENTION: REPEAT DNA MARKERS
; NUMBER OF SEQUENCES: 147
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Promega Corporation
; STREET: 2800 Woods Hollow Road
; CITY: Madison
; STATE: Wisconsin
; COUNTRY: U.S.A.
; ZIP: 53711-5399
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Mb
; COMPUTER: IBM compatible PC
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Word 97 (DOS text format)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/018,584A
; FILING DATE: 04-Feb-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Grady J. Frenchick
; REGISTRATION NUMBER: 29,018
; REFERENCE/DOCKET NUMBER: 16026.9180
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 257-3501
; TELEFAX: (608) 257-2275
; INFORMATION FOR SEQ ID NO: 124:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-09-018-584A-124

Query Match 100.0%; Score 25; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 4.1e-06;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTTCAGTGAGCGGAGATAGAGT 25
|||||
DB 1 GGTTCAGTGAGCGGAGATAGAGT 25

RESULT 2
 US-09-018-584A-32
 ; Sequence 32, Application US/09018584A
 ; Patent No. 6238863
 ; GENERAL INFORMATION:
 ; APPLICANT: Schumm, James W.
 ; TITLE OF INVENTION: MATERIALS AND METHODS FOR
 ; IDENTIFYING AND ANALYZING INTERMEDIATE TANDEM
 ; REPEAT DNA MARKERS
 ; NUMBER OF SEQUENCES: 147
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Promega Corporation
 ; STREET: 2800 Woods Hollow Road
 ; CITY: Madison
 ; STATE: Wisconsin
 ; COUNTRY: U.S.A.
 ; ZIP: 53711-5399
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Mb
 ; COMPUTER: IBM compatible PC
 ; OPERATING SYSTEM: Windows 95
 ; SOFTWARE: Word 97 (DOS text format)
 ; CURRENT APPLICATION NUMBER:
 ; APPLICATION NUMBER: US/09/018,584A
 ; FILING DATE: 04-Feb-1998
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Grady J. Frenchick
 ; REGISTRATION NUMBER: 29,018
 ; REFERENCE/DOCKET NUMBER: 16026.9180
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (608) 257-3501
 ; TELEFAX: (608) 257-2275
 ; INFORMATION FOR SEQ ID NO: 32:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1000 bp
 ; TYPE: Nucleic Acid
 ; STRANDEDNESS: Double
 ; TOPOLOGY: Circular
 ; MOLECULE TYPE: Genomic DNA
 ; HYPOTHETICAL: NO
 ; IMMEDIATE SOURCE:
 ; CLONE: S132
 ; POSITION IN GENOME:
 ; CHROMOSOME/SEGMENT: 22
 ; US-09-018-584A-32

Query Match 100.0%; Score 25; DB 4; Length 1000;
 Best Local Similarity 100.0%; Pred. No. 4e-06;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATAAGAGT 25
 ||||||||||||||||||||
 Db 441 GGTTCAGTGAGCCGAGATAAGAGT 465

RESULT 3
 US-09-288-143-57
 ; Sequence 57, Application US/09288143
 ; Patent No. 6433139
 ; GENERAL INFORMATION:
 ; APPLICANT: Brewer et al.
 ; TITLE OF INVENTION: 53 Human Secreted Proteins
 ; FILE REFERENCE: P2018P1
 ; CURRENT APPLICATION NUMBER: US/09/288,143
 ; CURRENT FILING DATE: 1999-04-08
 ; EARLIER APPLICATION NUMBER: PCT/US98/21142
 ; EARLIER FILING DATE: 1998-10-08
 ; EARLIER APPLICATION NUMBER: 60/061,463
 ; EARLIER FILING DATE: 1997-10-09
 ; EARLIER APPLICATION NUMBER: 60/061,529

; EARLIER FILING DATE: 1997-10-09
 ; EARLIER APPLICATION NUMBER: 60/071,498
 ; EARLIER FILING DATE: 1997-10-09
 ; EARLIER APPLICATION NUMBER: 60/061,527
 ; EARLIER FILING DATE: 1997-10-09
 ; EARLIER APPLICATION NUMBER: 60/061,536
 ; EARLIER FILING DATE: 1997-10-09
 ; EARLIER APPLICATION NUMBER: 60/061,532
 ; EARLIER FILING DATE: 1997-10-09
 ; NUMBER OF SEQ ID NOS: 219
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 57
 ; LENGTH: 764
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-288-143-57

Query Match 80.0%; Score 20; DB 4; Length 764;
 Best Local Similarity 100.0%; Pred. No. 0.0032;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATA 20
 ||||||||||||||||||||
 Db 674 GGTTCAGTGAGCCGAGATA 693

RESULT 4
 US-08-683-743-3
 ; Sequence 3, Application US/08683743
 ; Patent No. 5843697
 ; GENERAL INFORMATION:
 ; APPLICANT: Pestka, Sidney
 ; APPLICANT: Kotenko, Serguei
 ; TITLE OF INVENTION: CYTOKINE RECEPTOR SIGNAL TRANSDUCTION
 ; TITLE OF INVENTION: CHAIN
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: David A. Jackson, Esq.
 ; STREET: 411 Hackensack Ave, Continental Plaza, 4th
 ; STREET: Floor
 ; CITY: Hackensack
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07601
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/683,743
 ; FILING DATE: 17-JUL-1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Jackson Esq., David A.
 ; REGISTRATION NUMBER: 26,742
 ; REFERENCE/DOCKET NUMBER: 601-1-050
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-487-5800
 ; TELEFAX: 201-343-1684
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1875 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHETICAL: NO
 ; US-08-683-743-3

Query Match 80.0%; Score 20; DB 2; Length 1875;
 Best Local Similarity 100.0%; Pred. No. 0.0032;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGGCCGAGATA 20
|||||
Db 1551 GGTTCAGTGGCCGAGATA 1570

RESULT 5

US-08-406-030A-17
; Sequence 17, Application US/08406030A
; Patent No. 6270989
; GENERAL INFORMATION:
; APPLICANT: Treco, Douglas A.
; APPLICANT: Heartlein, Michael W.
; APPLICANT: Hauege, Brian M.
; APPLICANT: Selden, Richard F.
; TITLE OF INVENTION: Protein Production and Delivery
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/406,030A
; FILING DATE: 17-MAR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/243,391
; FILING DATE: 13-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/985,586
; FILING DATE: 03-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/911,533
; FILING DATE: 10-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,840
; FILING DATE: 05-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/789,188
; FILING DATE: 05-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11704
; FILING DATE: 02-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/09627
; FILING DATE: 05-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Grandhan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: TKT95-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4042 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-406-030A-17

Query Match 80.0%; Score 20; DB 4; Length 4042;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGGCCGAGATA 20
|||||
Db 2894 GGTTCAGTGGCCGAGATA 2913

RESULT 6

US-08-370-319C-12
; Sequence 12, Application US/08370319C
; Patent No. 5856091
; GENERAL INFORMATION:
; APPLICANT: Brichard, Vincent; Van Pel, Aline;
; APPLICANT: Traversari, Catia; Wlfel, Thomas; Coulie, Pierre;
; APPLICANT: Boon-Falleur, Thierry; De Plaen, Etienne
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID SEQUENCE CODING FOR A
; TITLE OF INVENTION: TUMOR REJECTION ANTIGEN PERCURSOR PROCESSED TO AT LEAST ONE
; TITLE OF INVENTION: REJECTION ANTIGEN PRESENTED BY HLA-A2
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,319C
; FILING DATE: 10-JANUARY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/272,351
; FILING DATE: 8-JULY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/032,978
; FILING DATE: 18-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5856091man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5377.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4129 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION: The sequence is preceded by an
; OTHER INFORMATION: unsequenced portion of from 4.7 to 5.3
; OTHER INFORMATION: kilobases
US-08-370-319C-12

Query Match 80.0%; Score 20; DB 2; Length 4129;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGGCCGAGATA 20
|||||
Db 3066 GGTTCAGTGGCCGAGATA 3085

RESULT 7

US-09-224-834-12
; Sequence 12, Application US/09224834
; Patent No. 6201111
; GENERAL INFORMATION:


```
; APPLICANT: Brichard, Vincent; Van Pel, Aline;
; APPLICANT: Traversari, Catia; W Ifel, Thomas; Coullie, Pierre;
; APPLICANT: Bcon-failleur, Thierry; De plaen, Etienne
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID SEQUENCE CODING FOR A
; TITLE OF INVENTION: TUMOR REJECTION ANTIGEN PERCURSOR PROCESSED TO AT LEAST ONE TU
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/224,834
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,319
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/032,978
; FILING DATE: 18-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 620111man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: IUD 5377.1
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4129 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION: The sequence is preceded by an
; OTHER INFORMATION: unsequenced portion of from 4.7 to 5.3
; OTHER INFORMATION: kilobases
; US-09-224-834-12

Query Match
Best Local Similarity 80.0%; Score 20; DB 4; Length 4129;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATA 20
Db 3066 GGTTCAGTGAGCCGAGATA 3085

RESULT 8
US-08-965-048-5
; Sequence 5, Application US/08965048
; Patent No. 6323244
; GENERAL INFORMATION:
; APPLICANT: Chen, Hong
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND
; FILE REFERENCE: 7853-093
; CURRENT APPLICATION NUMBER: US/08/965,048
; CURRENT FILING DATE: 1997-11-05
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 45716

; APPLICANT: Brichard, Vincent; Van Pel, Aline;
; APPLICANT: Traversari, Catia; W Ifel, Thomas; Coullie, Pierre;
; APPLICANT: Bcon-failleur, Thierry; De plaen, Etienne
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID SEQUENCE CODING FOR A
; TITLE OF INVENTION: TUMOR REJECTION ANTIGEN PERCURSOR PROCESSED TO AT LEAST ONE TU
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/224,834
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,319
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/032,978
; FILING DATE: 18-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 620111man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: IUD 5377.1
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4129 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION: The sequence is preceded by an
; OTHER INFORMATION: unsequenced portion of from 4.7 to 5.3
; OTHER INFORMATION: kilobases
; US-09-224-834-12

Query Match
Best Local Similarity 100.0%; Score 20; DB 4; Length 45716;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATA 20
Db 33645 GGTTCAGTGAGCCGAGATA 33664

RESULT 9
US-08-965-048-6
; Sequence 6, Application US/08965048
; Patent No. 6323244
; GENERAL INFORMATION:
; APPLICANT: Chen, Hong
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND
; FILE REFERENCE: 7853-093
; CURRENT APPLICATION NUMBER: US/08/965,048
; CURRENT FILING DATE: 1997-11-05
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 45989
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-08-965-048-6

Query Match
Best Local Similarity 100.0%; Score 20; DB 4; Length 45989;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATA 20
Db 33645 GGTTCAGTGAGCCGAGATA 33664

RESULT 10
US-08-133-629-2
; Sequence 2, Application US/08133629
; Patent No. 5597694
; GENERAL INFORMATION:
; APPLICANT: Munroe, David J.
; APPLICANT: Housman, David E.
; TITLE OF INVENTION: AMPLIFICATION OF NUCLEIC ACIDS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/133,629
; FILING DATE: 07-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Greer, Helen
; REGISTRATION NUMBER: 36,816
; REFERENCE/DOCKET NUMBER: M0828/7001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-08-965-048-5

Query Match
Best Local Similarity 80.0%; Score 20; DB 4; Length 45716;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATA 20
Db 33531 GGTTCAGTGAGCCGAGATA 33550

RESULT 9
US-08-965-048-6
; Sequence 6, Application US/08965048
; Patent No. 6323244
; GENERAL INFORMATION:
; APPLICANT: Chen, Hong
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND
; FILE REFERENCE: 7853-093
; CURRENT APPLICATION NUMBER: US/08/965,048
; CURRENT FILING DATE: 1997-11-05
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 45989
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-08-965-048-6

Query Match
Best Local Similarity 100.0%; Score 20; DB 4; Length 45989;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATA 20
Db 33645 GGTTCAGTGAGCCGAGATA 33664

RESULT 10
US-08-133-629-2
; Sequence 2, Application US/08133629
; Patent No. 5597694
; GENERAL INFORMATION:
; APPLICANT: Munroe, David J.
; APPLICANT: Housman, David E.
; TITLE OF INVENTION: AMPLIFICATION OF NUCLEIC ACIDS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/133,629
; FILING DATE: 07-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Greer, Helen
; REGISTRATION NUMBER: 36,816
; REFERENCE/DOCKET NUMBER: M0828/7001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
```

```

; TELEFAX: 617-720-2441
; TELEX: 92-1742 EZEKIEL
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-133-629-2

Query Match 76.0%; Score 19; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
Db 3 GGTTCAGTGAGCCGAGAT 21
|||||

RESULT 11
US-08-687-080-93/c
; Sequence 93, Application US/08687080
; Patent No. 5965427
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/687,080
; FILING DATE: 17-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/592,126
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0111.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 93:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 239 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: 5' END OF INTRON 16 OF RAD50 GENOMIC
; INDIVIDUAL ISOLATE: SEQUENCE
US-08-687-080-93

Query Match 76.0%; Score 19; DB 2; Length 239;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
|||||
```

```

Db 128 GGTTCAGTGAGCCGAGAT 110

RESULT 12
US-08-849-701-1
; Sequence 1, Application US/08849701
; Patent No. 5922544
; GENERAL INFORMATION:
; APPLICANT: Miyai, Kiyoshi
; APPLICANT: Naitch, Tsutomu
; APPLICANT: Yonekawa, Toshihiro
; TITLE OF INVENTION: Method of Cell Detection
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,701
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP95/02734
; FILING DATE: 27-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: EIKEN1.001APC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714-760-0404
; TELEFAX: 714-760-9502
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 265 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Other
; IMMEDIATE SOURCE:
; CLONE: Alu sequence BLUR8
US-08-849-701-1

Query Match 76.0%; Score 19; DB 2; Length 265;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
|||||

Db 194 GGTTCAGTGAGCCGAGAT 212

RESULT 13
US-08-481-658B-61/c
; Sequence 61, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
```

; TITLE OF INVENTION: BACTERIUM PRODUCING L-GLUTAMIC ACID AND METHOD FOR PRODUCING L-GI
; FILE REFERENCE: 210213US0
; CURRENT APPLICATION NUMBER: US/09/895,382
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: JP 2000-204256
; PRIOR FILING DATE: 2000-07-05
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic DNA
US-09-895-382-6

Query Match 52.0%; Score 13; DB 10; Length 24;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TGCAGGAACACG 15
|||||
Db 15 TGCAGGAACACG 3

RESULT 21
US-09-946-807-1310/c
; Sequence 1310, Application US/09946807
; Patent No. US20020165144A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn
; APPLICANT: Steinthorsdottir, Valgerdur
; APPLICANT: Gulcher, Jeffrey R.
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345.2004-001
; CURRENT APPLICATION NUMBER: US/09/946,807
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US/09/795,668
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/515,716
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1310
; LENGTH: 61
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-946-807-1310

Query Match 52.0%; Score 13; DB 9; Length 61;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 AACCAAGAAATTTA 22
|||||
Db 52 AACCAAGAAATTTA 40

RESULT 22
US-09-795-668-1310/c
; Sequence 1310, Application US/09795668
; Patent No. US20020045577A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn
; APPLICANT: Steinthorsdottir, Valgerdur
; APPLICANT: Gulcher, Jeffrey R.
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345.2004-001
; CURRENT APPLICATION NUMBER: US/09/795,668
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/515,716
; PRIOR FILING DATE: 2000-02-28

; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1310
; LENGTH: 61
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-795-668-1310

Query Match 52.0%; Score 13; DB 10; Length 61;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 AACCAAGAAATTTA 22
|||||
Db 52 AACCAAGAAATTTA 40

RESULT 23
US-09-795-686-1310/c
; Sequence 1310, Application US/09795686
; Patent No. US2002009454A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn
; APPLICANT: Steinthorsdottir, Valgerdur
; APPLICANT: Gulcher, Jeffrey R.
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345.2005-001
; CURRENT APPLICATION NUMBER: US/09/795,686
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/515,715
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1310
; LENGTH: 61
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-795-686-1310

Query Match 52.0%; Score 13; DB 10; Length 61;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 AACCAAGAAATTTA 22
|||||
Db 52 AACCAAGAAATTTA 40

RESULT 24
US-09-878-574-11569/c
; Sequence 11569, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Thompson, Michael J.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 11569
; LENGTH: 157
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701064613H1
US-09-878-574-11569

Query Match 52.0%; Score 13; DB 10; Length 157;
Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTGCCAGGACCA 14
Db 71 GTGCCAGGACCA 59

RESULT 25
US-09-960-352-2301/C
; Sequence 2301, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Neungbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 2301
; LENGTH: 209
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 10-LIB34-059-01-El-C9
US-09-960-352-2301

Query Match 52.0%; Score 13; DB 10; Length 209;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AGGACCCAGGAAT 19
Db 189 AGGACCCAGGAAT 177

RESULT 26
US-09-878-574-14842/C
; Sequence 14842, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 14842
; LENGTH: 212
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701069323H1
US-09-878-574-14842

Query Match 52.0%; Score 13; DB 10; Length 212;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGACCC 13
Db 196 TGTGCCAGGACCC 184

RESULT 27
US-09-864-761-21959
; Sequence 21959, Application US/09864761

; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/006666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21959
; LENGTH: 223
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004585.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.1
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5
; OTHER INFORMATION: SWISSPROT HIT: Q14031, EVALUATE 2.70e-01
; OTHER INFORMATION: EST_HUMAN HIT: A1361780.1, EVALUATE 5.10e-02
; OTHER INFORMATION: NT HIT: U24393.1, EVALUATE 3.40e-02
US-09-864-761-21959

Query Match 52.0%; Score 13; DB 10; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGACCC 13

```
Db      15  TGTGCCAGGAACC 27
|||||
RESULT 28
US-09-878-574-6327/c
; Sequence 6327, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 6327
; LENGTH: 247
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701098192H1
US-09-878-574-6327

Query Match      52.0%; Score 13; DB 10; Length 247;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2  GTGCCAGGAACCA 14
|||||
Db      32  GTGCCAGGAACCA 20

RESULT 29
US-09-878-574-10539/c
; Sequence 10539, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 10539
; LENGTH: 261
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 700966787H1
US-09-878-574-10539

Query Match      52.0%; Score 13; DB 10; Length 261;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2  GTGCCAGGAACCA 14
|||||
Db      187 GTGCCAGGAACCA 175

RESULT 30
US-09-983-965-117
; Sequence 117, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 117
; LENGTH: 273
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; OTHER INFORMATION: Clone ID: 14-BOYMS1-012-Q1-EL-D9
US-09-983-965-117

Query Match      52.0%; Score 13; DB 10; Length 273;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6  CAGGAACCAAGAA 18
|||||
Db      7  CAGGAACCAAGAA 19

RESULT 31
US-09-783-590-11124/c
; Sequence 11124, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11124
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (9)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (30)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (65)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (112)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (121)
```

OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (134)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (151)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (159)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (166)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (186)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (194)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (205)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (238)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (240)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (242)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (256)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (281)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (290)
OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-11124

Query Match 52.0%; Score 13; DB 10; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 AGGAACCCAGAAAT 19
|||||
Db 276 AGGAACCCAGAAAT 264

RESULT 32
US-09-960-352-2711/c
; Sequence 2711, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 2711
; LENGTH: 382
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 12-LIB34-005-Q1-E1-C7
US-09-960-352-2711

Query Match 52.0%; Score 13; DB 10; Length 382;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 AGGAACCCAGAAAT 19
|||||
Db 229 AGGAACCCAGAAAT 217
RESULT 33
US-09-960-352-4653/c
; Sequence 4653, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4653
; LENGTH: 395
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 20-LIB34-012-Q1-E1-E7
US-09-960-352-4653

Query Match 52.0%; Score 13; DB 10; Length 395;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 AGGAACCCAGAAAT 19
|||||
Db 349 AGGAACCCAGAAAT 337

RESULT 34
US-09-960-352-7956/c
; Sequence 7956, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 7956
; LENGTH: 404
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 34-LIB34-005-Q1-E1-A6
US-09-960-352-7956

Query Match 52.0%; Score 13; DB 10; Length 404;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 AGGAACCCAGAAAT 19
|||||
Db 22 AGGAACCCAGAAAT 10

RESULT 35
US-09-864-761-6412/c

```

; Sequence 6412, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Acomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.5
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 6412
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005924.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.4
; US-09-864-761-6412

```

```

Query Match 52.0%; Score 13; DB 10; Length 405;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 GTGCCAGGAACCA 14
    |||||
DB 161 GTGCCAGGAACCA 149

```

```

RESULT 36
US-09-960-352-2744/c
; Sequence 2744, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; NUMBER OF SEQ ID NOS: 2001-09-24
; SEQ ID NO 2744
; LENGTH: 416
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 12-LIB34-031-QL-E1-C7
; US-09-960-352-2744

Query Match 52.0%; Score 13; DB 10; Length 416;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AGGACCAAGAAAT 19
    |||||
DB 294 AGGACCAAGAAAT 282

RESULT 37
US-09-783-590-6944
; Sequence 6944, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16, 2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6944
; LENGTH: 475
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (51)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (204)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (297)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (302)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature

```

LOCATION: (322)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (336)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (339)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (352)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (356)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (370)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (371)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (392)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (396)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (412)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (428)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (441)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (454)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (459)
OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-6944

Query Match 52.0%; Score 13; DB 10; Length 475;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTGCCAGGAAC 13
Db 92 TGTGCCAGGAAC 104

RESULT 38
US-09-878-574-4307/c
Sequence 4307, Application US/09878574
Patent No. US20020110549A1
GENERAL INFORMATION:
APPLICANT: Byrum, Joseph R.
APPLICANT: La Rosa, Thomas J.
APPLICANT: Thompson, Michael D.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(15401)B
CURRENT APPLICATION NUMBER: US/09/878,574
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 09/333,535
PRIOR FILING DATE: 1999-06-14
NUMBER OF SEQ ID NOS: 15775
SEQ ID NO 4307
LENGTH: 479
TYPE: DNA
ORGANISM: Glycine max
FEATURE:

NAME/KEY: unsure
LOCATION: (1)..(479)
OTHER INFORMATION: unsure at all n locations
OTHER INFORMATION: Clone ID: LIB3028-014-01-B1-E10
US-09-878-574-4307

Query Match 52.0%; Score 13; DB 10; Length 479;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 CAGGACACAGAAA 18
Db 289 CAGGACACAGAAA 277

RESULT 39
US-09-864-761-6095
Sequence 6095, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/006666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/006657
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/006654
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/006669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/006665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/006668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/006663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/006662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/006661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/006670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 6095
LENGTH: 481
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004585.1


```

; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5
US-09-864-761-6095

Query Match      52.0%; Score 13; DB 10; Length 481;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTGCCAGGAACC 13
Db 286 TGTGCCAGGAACC 298

RESULT 40
US-09-783-590-11922
; Sequence 11922, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11922
; LENGTH: 489
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (73)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (88)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (100)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (275)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (305)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (333)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (351)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (397)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (422)

```

```

; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (445)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (454)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (460)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (466)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (473)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-11922

```

```

Query Match      52.0%; Score 13; DB 10; Length 489;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 8 GGAACCAAGAAATT 20
Db 208 GGAACCAAGAAATT 220

```

```

RESULT 41
US-09-833-790-163
; Sequence 163, Application US/09833790
; Patent No. US20020068288A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Secrist, Heather
; APPLICANT: Mohamath, Raodon
; APPLICANT: Indrias, Carol Y.
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121-512
; CURRENT APPLICATION NUMBER: US/09/833,790
; CURRENT FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 163
; LENGTH: 501
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(501)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-790-163

```

```

Query Match      52.0%; Score 13; DB 10; Length 501;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 13 CAGAAATTACAG 25
Db 257 CAGAAATTACAG 269

```

```

RESULT 42
US-09-783-590-5680/c
; Sequence 5680, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.

```

```

; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5680
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (11)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (109)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (119)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (214)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (282)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (306)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (314)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (320)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (322)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (329)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (334)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (346)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (364)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (394)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (396)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (399)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (406)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (411)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (424)
; OTHER INFORMATION: n equals a,t,g, or c

```

```

; NAME/KEY: misc feature
; LOCATION: (425)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (430)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (434)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (435)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (439)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (441)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (444)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (451)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (453)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (467)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (471)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (474)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (478)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (479)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (481)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (484)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (496)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-783-590-5680

```

```

Query Match 52.0%; Score 13; DB 10; Length 506;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 1 TGTGCCAGGACC 13
    |||||
Db 250 TGTGCCAGGACC 238

```

```

RESULT 43
US-09-974-300-7386
; Sequence 7386, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300

```

; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7386
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Bacillus clausii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(531)
; OTHER INFORMATION: n = A,T,C or G
US-09-974-300-7386

Query Match 52.0%; Score 13; DB 10; Length 531;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 GGAACCGAATTT 20
Db 305 GGAACCGAATTT 317
|||||

RESULT 44

US-10-062-254-119
; Sequence 119, Application US/10062254
; Patent No. US20020138882A1
; GENERAL INFORMATION:
; APPLICANT: Caboon, Edgar B
; APPLICANT: Caboon, Rebecca E
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Fang, Yiwen
; APPLICANT: Hantke, Sabine S.
; APPLICANT: Lee, Jian-Ming
; APPLICANT: Li, Zhongsen
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Morgante, Michele
; APPLICANT: Niu, Xiping
; APPLICANT: Odell, Joan
; APPLICANT: Rafalski, Antoni
; APPLICANT: Sakai, Hajime
; APPLICANT: Zheng, Peizhong
; APPLICANT: Zhu, Qun
; TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/062,254
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 09/630,346
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/146511
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 60/156006
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/156899
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 60/157287
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 60/169767
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/171054
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/172958
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/171515
; PRIOR FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: 60/173535
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 375
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 119

; LENGTH: 531
; TYPE: DNA
; ORGANISM: Vernonia mespilifolia
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (422)
; NAME/KEY: unsure
; LOCATION: (528)
US-10-062-254-119

Query Match 52.0%; Score 13; DB 12; Length 531;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 CCAGAAATTACA 24
Db 363 CCAGAAATTACA 375
|||||

RESULT 45

US-09-834-975-60
; Sequence 60, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Huffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MRI-016B
; CURRENT APPLICATION NUMBER: US/09/834,975
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(535)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-60

Query Match 52.0%; Score 13; DB 10; Length 535;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 CCAGAAATTACA 24
Db 370 CCAGAAATTACA 382
|||||

Search completed: November 24, 2002, 16:56:07
Job time : 1156.5 secs

QY 1 GGTTCAGTGAGCCGAGAT 19
Db 75 GGTTCAGTGAGCCGAGAT 57

RESULT 14
US-08-477-504A-61/c
; Sequence 61, Application US/08477504A
; Patent No. 5972353
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,504A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 294 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-481-658B-61

Query Match 76.0%; Score 19; DB 2; Length 294;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
Db 75 GGTTCAGTGAGCCGAGAT 57

RESULT 15
US-08-486-756A-61/c
; Sequence 61, Application US/08486756A
; Patent No. 5981711
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,756A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 294 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-486-756A-61

QY 1 GGTTCAGTGAGCCGAGAT 19
Db 75 GGTTCAGTGAGCCGAGAT 57

RESULT 15
US-08-486-756A-61/c
; Sequence 61, Application US/08486756A
; Patent No. 5981711
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,756A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 294 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-486-756A-61

Query Match 76.0%; Score 19; DB 2; Length 294;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
Db 75 GGTTCAGTGAGCCGAGAT 57

RESULT 15
US-08-486-756A-61/c
; Sequence 61, Application US/08486756A
; Patent No. 5981711
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,756A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 294 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-486-756A-61

Query Match 76.0%; Score 19; DB 2; Length 294;
 Best Local Similarity 100.0%; Pred. No. 0.012;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGCAGCCGAGAT 19
 |||||
 Db 75 GGTTCAGTGCAGCCGAGAT 57

RESULT 16

US-08-485-862B-61/c
 ; Sequence 61, Application US/08485862B
 ; Patent No. 5989838

GENERAL INFORMATION:
 ; APPLICANT: Zavada, Jan
 ; APPLICANT: Pastorekova, Silvia
 ; APPLICANT: Pastorek, Jaromir
 ; TITLE OF INVENTION: MN Gene and Protein
 ; NUMBER OF SEQUENCES: 86
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Leona L. Lauder
 ; STREET: 6 Mariposa Court
 ; CITY: Tiburon
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94920

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/485,862B
 FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/477,504
 FILING DATE: 07-JUN-1995
 APPLICATION NUMBER: US 08/260,190
 FILING DATE: 15-JUN-1994

ATTORNEY/AGENT INFORMATION:
 NAME: Lauder, Leona L.
 REGISTRATION NUMBER: 30,863
 REFERENCE/DOCKET NUMBER: D-0021.3D
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-435-2034
 TELEFAX: 415-435-0727

INFORMATION FOR SEQ ID NO: 61:

SEQUENCE CHARACTERISTICS:
 LENGTH: 294 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-485-862B-61

Query Match 76.0%; Score 19; DB 2; Length 294;
 Best Local Similarity 100.0%; Pred. No. 0.012;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGCAGCCGAGAT 19
 |||||
 Db 75 GGTTCAGTGCAGCCGAGAT 57

RESULT 17

US-08-787-739-61/c
 ; Sequence 61, Application US/08787739
 ; Patent No. 6027887

GENERAL INFORMATION:
 ; APPLICANT: Zavada, Jan

APPLICANT: Pastorekova, Silvia
 APPLICANT: Pastorek, Jaromir
 TITLE OF INVENTION: MN Gene and Protein
 NUMBER OF SEQUENCES: 96
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Leona L. Lauder
 STREET: 369 Pine Street, Suite 610
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/787,739

FILING DATE: 24-JAN-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/485,049
 FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/486,756
 FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/477,504
 FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/481,658
 FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/485,862
 FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/485,863
 FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/487,077
 FILING DATE: 07-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Lauder, Leona L.

REGISTRATION NUMBER: 30,863

REFERENCE/DOCKET NUMBER: D-0021.4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-981-0332

TELEFAX: 415-981-0332

INFORMATION FOR SEQ ID NO: 61:

SEQUENCE CHARACTERISTICS:

LENGTH: 294 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-787-739-61

Query Match 76.0%; Score 19; DB 3; Length 294;
 Best Local Similarity 100.0%; Pred. No. 0.012;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGCAGCCGAGAT 19
 |||||
 Db 75 GGTTCAGTGCAGCCGAGAT 57

RESULT 18

US-08-487-077A-61/c
 ; Sequence 61, Application US/08487077A
 ; Patent No. 6069242

GENERAL INFORMATION:
 ; APPLICANT: Zavada, Jan

```

; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,077A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION INFORMATION:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; NAME: Lauder, Leona L.
; ATTORNEY/AGENT INFORMATION:
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3H
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-0727
; FAX: 415-435-2034
; INFORMATION FOR SEQ ID NO: 61:
; LENGTH: 294 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-487-077A-61

Query Match 76.0%; Score 19; DB 3; Length 294;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
Db 75 GGTTCAGTGAGCCGAGAT 57

RESULT 19
US-08-485-863A-61/c
; Sequence 61, Application US/08485863A
; Patent No. 6093546
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,077A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION INFORMATION:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; NAME: Lauder, Leona L.
; ATTORNEY/AGENT INFORMATION:
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; FAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 61:
; LENGTH: 294 base pairs
; TYPE: nucleic acid
; US-08-487-077A-61
```

```

; APPLICATION NUMBER: US/08/485,863A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; FAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 61:
; LENGTH: 294 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-485-863A-61

Query Match 76.0%; Score 19; DB 3; Length 294;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
Db 75 GGTTCAGTGAGCCGAGAT 57

RESULT 20
US-08-485-049D-61/c
; Sequence 61, Application US/08485049D
; Patent No. 6204370
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,049D
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; FAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 61:
; LENGTH: 294 base pairs
; TYPE: nucleic acid
; US-08-485-049D-61
```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-485-049D-61

```

```

Query Match          76.0%; Score 19; DB 4; Length 294;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGTTCAGTGAGCCGAGAT 19
    |||||||
Db 75 GGTTCAGTGAGCCGAGAT 57

```

RESULT 21

```

US-09-178-115-61/c
; Sequence 61, Application US/09178115
; Patent No. 6297041
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; FILE REFERENCE: D-0021.5A
; CURRENT APPLICATION NUMBER: US/09/178,115
; EARLIER FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 09/177,776
; EARLIER FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 08/787,739
; EARLIER FILING DATE: 1997-01-24
; EARLIER APPLICATION NUMBER: 08/485,049
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/486,756
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/477,504
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/481,658
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,862
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,863
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/487,077
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/260,190
; EARLIER FILING DATE: 1994-06-15
; EARLIER APPLICATION NUMBER: 08/177,093
; EARLIER FILING DATE: 1993-12-30
; EARLIER APPLICATION NUMBER: 07/964,589
; EARLIER FILING DATE: 1992-10-21
; EARLIER APPLICATION NUMBER: PV-709-92
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61
; LENGTH: 294
; TYPE: DNA
; ORGANISM: HUMAN
US-09-178-115-61

```

```

Query Match          76.0%; Score 19; DB 4; Length 294;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGTTCAGTGAGCCGAGAT 19
    |||||||
Db 75 GGTTCAGTGAGCCGAGAT 57

```

RESULT 22

US-09-177-776-61/c

```

; Sequence 61, Application US/09177776A
; Patent No. 6297051
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; FILE REFERENCE: D-0021.5A
; CURRENT APPLICATION NUMBER: US/09/177,776A
; EARLIER FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 08/787,739
; EARLIER FILING DATE: 1997-01-24
; EARLIER APPLICATION NUMBER: 08/485,049
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/486,756
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/477,504
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/481,658
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,862
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,863
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/487,077
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/260,190
; EARLIER FILING DATE: 1994-06-15
; EARLIER APPLICATION NUMBER: 08/177,093
; EARLIER FILING DATE: 1993-12-30
; EARLIER APPLICATION NUMBER: 07/964,589
; EARLIER FILING DATE: 1992-10-21
; EARLIER APPLICATION NUMBER: PV-709-92
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61
; LENGTH: 294
; TYPE: DNA
; ORGANISM: HUMAN
US-09-177-776-61

```

```

Query Match          76.0%; Score 19; DB 4; Length 294;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGTTCAGTGAGCCGAGAT 19
    |||||||
Db 75 GGTTCAGTGAGCCGAGAT 57

```

RESULT 23

```

US-08-849-701-3
; Sequence 3, Application US/08849701
; Patent No. 5922544
; GENERAL INFORMATION:
; APPLICANT: Miyai, Kiyoshi
; APPLICANT: Naitoh, Tsutomu
; APPLICANT: Yonekawa, Yoshihiro
; TITLE OF INVENTION: Method of Cell Detection
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5

```

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/849,701
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/JP95/02734
;; FILING DATE: 27-DEC-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Altman, Daniel E
;; REGISTRATION NUMBER: 34,115
;; REFERENCE/DOCKET NUMBER: EIKEN1.001APC
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 714-760-0404
;; TELEFAX: 714-760-9502
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 302 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; CLONE: Alu sequence BLUR2
US-08-849-701-3

Query Match 76.0%; Score 19; DB 2; Length 302;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
|||||
DB 217 GGTTCAGTGAGCCGAGAT 235

RESULT 24

US-09-385-982-17
;; Sequence 17, Application US/09385982
;; Patent No. 6262334
;; GENERAL INFORMATION:
;; APPLICANT: ENDEGE, WILSON O., ET AL.
;; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
;; FILE REFERENCE: CCNDA-260XX
;; CURRENT APPLICATION NUMBER: US/09/385,982
;; CURRENT FILING DATE: 1999-08-30
;; EARLIER APPLICATION NUMBER: 09/328,111
;; EARLIER FILING DATE: 1999-06-08
;; EARLIER APPLICATION NUMBER: 60/117,393
;; EARLIER FILING DATE: 1999-01-27
;; EARLIER APPLICATION NUMBER: 60/098,639
;; EARLIER FILING DATE: 1998-08-31
;; NUMBER OF SEQ ID NOS: 544
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 17
;; LENGTH: 336
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(336)
;; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-17

Query Match 76.0%; Score 19; DB 4; Length 336;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
|||||
DB 240 GGTTCAGTGAGCCGAGAT 258

RESULT 25

US-09-385-982-135/c
;; Sequence 135, Application US/09385982
;; Patent No. 6262334
;; GENERAL INFORMATION:
;; APPLICANT: ENDEGE, WILSON O., ET AL.
;; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
;; FILE REFERENCE: CCNDA-260XX
;; CURRENT APPLICATION NUMBER: US/09/385,982
;; CURRENT FILING DATE: 1999-08-30
;; EARLIER APPLICATION NUMBER: 09/328,111
;; EARLIER FILING DATE: 1999-06-08
;; EARLIER APPLICATION NUMBER: 60/117,393
;; EARLIER FILING DATE: 1999-01-27
;; EARLIER APPLICATION NUMBER: 60/098,639
;; EARLIER FILING DATE: 1998-08-31
;; NUMBER OF SEQ ID NOS: 544
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 135
;; LENGTH: 374
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(374)
;; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-135

Query Match 76.0%; Score 19; DB 4; Length 374;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
|||||
DB 95 GGTTCAGTGAGCCGAGAT 77

RESULT 26

US-09-157-177-134/c
;; Sequence 134, Application US/09157177
;; Patent No. 6090558
;; GENERAL INFORMATION:
;; APPLICANT: Butler, John M.
;; APPLICANT: Li, Jia
;; APPLICANT: Monforte, Joseph A.
;; APPLICANT: Becker, Christopher H.
;; TITLE OF INVENTION: DNA TYPING BY MASS SPECTROMETRY WITH POLYMORPHIC DNA
;; FILE REFERENCE: GSTR:017/GETR017P
;; CURRENT APPLICATION NUMBER: US/09/157,177
;; CURRENT FILING DATE: 1998-09-18
;; EARLIER APPLICATION NUMBER: 60/059,415
;; EARLIER FILING DATE: 1997-09-19
;; NUMBER OF SEQ ID NOS: 135
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 134
;; LENGTH: 379
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-157-177-134

Query Match 76.0%; Score 19; DB 3; Length 379;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
|||||
DB 355 GGTTCAGTGAGCCGAGAT 337

RESULT 27

US-09-227-357-92
;; Sequence 92, Application US/09227357

Patent No. 6342581
GENERAL INFORMATION:
APPLICANT: Fischer et al.
TITLE OF INVENTION: 123 Human Secreted Proteins
FILE REFERENCE: P2010P1
CURRENT APPLICATION NUMBER: US/09/227,357
CURRENT FILING DATE: 1999-01-08
EARLIER APPLICATION NUMBER: PCT/US98/13684
EARLIER FILING DATE: 1998-07-07
EARLIER APPLICATION NUMBER: 60/051,926
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,793
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,925
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,929
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,803
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,732
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,931
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,932
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,916
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,930
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,918
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,920
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,733
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,795
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,919
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,928
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/055,722
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,723
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,948
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,964
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,684
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,785
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
EARLIER FILING DATE: 1997-09-12

NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 92
LENGTH: 577
TYPE: DNA
ORGANISM: Homo sapiens
US-09-227-357-92

Query Match 76.0%; Score 19; DB 4; Length 577;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGAT 19
|||||
Db 489 GGTTCAGTGCAGCCGAGAT 507

RESULT 28
US-09-385-982-291/c
Sequence 291, Application US/09385982
Patent No. 6262334
GENERAL INFORMATION:
APPLICANT: ENDEGE, WILSON O., ET AL.
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILE REFERENCE: CCDNA-260XX
CURRENT APPLICATION NUMBER: US/09/385,982
CURRENT FILING DATE: 1999-08-30
EARLIER APPLICATION NUMBER: 09/328,111
EARLIER FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: 60/117,393
EARLIER FILING DATE: 1999-01-27
EARLIER APPLICATION NUMBER: 60/098,639
EARLIER FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 544
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 291
LENGTH: 609
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(609)
OTHER INFORMATION: n = A,T,C or G
US-09-385-982-291

Query Match 76.0%; Score 19; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGAT 19
|||||
Db 30 GGTTCAGTGCAGCCGAGAT 12

RESULT 29
US-09-385-982-218/c
Sequence 218, Application US/09385982
Patent No. 6262334
GENERAL INFORMATION:
APPLICANT: ENDEGE, WILSON O., ET AL.
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILE REFERENCE: CCDNA-260XX
CURRENT APPLICATION NUMBER: US/09/385,982
CURRENT FILING DATE: 1999-08-30
EARLIER APPLICATION NUMBER: 09/328,111
EARLIER FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: 60/117,393
EARLIER FILING DATE: 1999-01-27
EARLIER APPLICATION NUMBER: 60/098,639
EARLIER FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 544

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 218
; LENGTH: 618
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(618)
; OTHER INFORMATION: n = A,T,C or G
US-09-383-982-218

Query Match 76.0%; Score 19; DB 4; Length 618;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
|||||
Db 94 GGTTCAGTGAGCCGAGAT 76

RESULT 30

US-09-535-008-49/c
; Sequence 49, Application US/09535008
; Patent No. 6465629

; GENERAL INFORMATION:
; APPLICANT: Wong, Alexander K.C.

; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.-F.

; TITLE OF INVENTION: BRG1 IS A TUMOR SUPPRESSOR THAT IS MUTATED IN PROSTATE
; FILE REFERENCE: 2318-259
; CURRENT APPLICATION NUMBER: US/09/535,008

; CURRENT FILING DATE: 2000-03-23
; EARLIER APPLICATION NUMBER: U.S. 60/125,806

; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 77

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49

; LENGTH: 649
; TYPE: DNA

; ORGANISM: Homo sapiens
US-09-535-008-49

Query Match 76.0%; Score 19; DB 4; Length 649;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
|||||
Db 584 GGTTCAGTGAGCCGAGAT 566

RESULT 31

US-09-105-542A-14
; Sequence 14, Application US/09105542A
; Patent No. 6323329

; GENERAL INFORMATION:
; APPLICANT: Bullerdick, John

; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES OF GENES OF THE
; FILE REFERENCE: BOEHM4.001C1P

; CURRENT APPLICATION NUMBER: US/09/105,542A
; CURRENT FILING DATE: 1998-06-26

; PRIOR APPLICATION NUMBER: PCT/DE96/02494
; PRIOR FILING DATE: 1996-12-20

; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 14
; LENGTH: 689

; TYPE: DNA
; ORGANISM: Homo Sapien

US-09-105-542A-14

Query Match 76.0%; Score 19; DB 4; Length 689;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
|||||
Db 557 GGTTCAGTGAGCCGAGAT 575

RESULT 32

US-09-740-235-16/c

; Sequence 16, Application US/09740235
; Patent No. 6448041

; GENERAL INFORMATION:
; APPLICANT: Wolven, Amy K.

; APPLICANT: Krasnow, Randi E.
; APPLICANT: Warren, Bridget A.

; APPLICANT: Baughn, Mariah R.
; FILE REFERENCE: PC-0031 CIP

; TITLE OF INVENTION: COLON CANCER MARKER

; CURRENT APPLICATION NUMBER: US/09/740,235
; CURRENT FILING DATE: 2000-12-18

; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PERL Program

; SEQ ID NO 16
; LENGTH: 696

; TYPE: DNA
; ORGANISM: Homo sapiens

; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6448041 7984383H1

US-09-740-235-16

Query Match 76.0%; Score 19; DB 4; Length 696;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
|||||
Db 278 GGTTCAGTGAGCCGAGAT 260

RESULT 33

US-09-227-357-108

; Sequence 108, Application US/09227357
; Patent No. 6342581

; GENERAL INFORMATION:
; APPLICANT: Fischer et al.

; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1

; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08

; EARLIER APPLICATION NUMBER: PCI/US98/13684
; EARLIER FILING DATE: 1998-07-07

; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08

; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08

; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08

; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08

; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08

; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08

; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08

; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08

; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08

; EARLIER APPLICATION NUMBER: 60/051,930

EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,918
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,920
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,733
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,795
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,919
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,928
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/055,722
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,723
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,948
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,684
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,785
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
EARLIER FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 108
LENGTH: 775
TYPE: DNA
ORGANISM: Homo sapiens
US-09-227-357-108

Query Match 76.0%; Score 19; DB 4; Length 775;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTGCACTGAGCCGAGAT 19
Db 679 GGTTGCACTGAGCCGAGAT 697
|||||

RESULT 34
US-08-394-152A-42/C
Sequence 42, Application US/08394152A
Patent No. 5935818
GENERAL INFORMATION:
APPLICANT: Israeli, Ron S.
APPLICANT: Heston, Warren D.W.
APPLICANT: Fair, William R.
TITLE OF INVENTION: PROSTATE-SPECIFIC MEMBRANE ANTIGEN AND
NUMBER OF SEQUENCES: 48

CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM 330 466 DX2
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/394,152A
FILING DATE: 24-FEB-95
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41426-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 789 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo Sapien
TISSUE TYPE: Carcinoma
IMMEDIATE SOURCE:
CLONE: Prostate Specific Membrane Antigen
US-08-394-152A-42

Query Match 76.0%; Score 19; DB 2; Length 789;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTGCACTGAGCCGAGAT 19
Db 487 GGTTGCACTGAGCCGAGAT 469
|||||

RESULT 35
US-09-342-681C-7
Sequence 7, Application US/09342681C
Patent No. 6355782
GENERAL INFORMATION:
APPLICANT: Zonana et al.
TITLE OF INVENTION: Hypohydrotic ectodermal dysplasia genes and proteins
FILE REFERENCE: 52978
CURRENT APPLICATION NUMBER: US/09/342,681C
CURRENT FILING DATE: 1999-06-29
PRIOR APPLICATION NUMBER: 60/092,279
PRIOR FILING DATE: 1998-07-09
PRIOR APPLICATION NUMBER: 60/112,366
PRIOR FILING DATE: 1998-12-15
NUMBER OF SEQ ID NOS: 123
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 821
TYPE: DNA
ORGANISM: Homo sapiens
US-09-342-681C-7

Query Match 76.0%; Score 19; DB 4; Length 821;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGCAGCGAGAT 19
Db 192 GGTTCAGTGCAGCGAGAT 210
|||||
RESULT 36
US-08-486-148B-1/c
; Sequence 1, Application US/08486148B
; Patent No. 5840555
; GENERAL INFORMATION:
; APPLICANT: Oshima, Robert G.
; APPLICANT: Meznarov, Nikolay S.
; TITLE OF INVENTION: Transcriptional Regulatory Regions
; TITLE OF INVENTION: Derived from the K18 Gene.
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,148B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/146,486
; FILING DATE: 29-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1698
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-486-148B-1
Query Match 76.0%; Score 19; DB 2; Length 825;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGTTCAGTGCAGCGAGAT 19
Db 522 GGTTCAGTGCAGCGAGAT 504
|||||
RESULT 37
PCT-US93-06251-80
; Sequence 80, Application PC/TUS9306251
; GENERAL INFORMATION:
; APPLICANT: Wickstrom, Eric and Rife, Jason P.
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City

STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06251
FILING DATE: 19930630
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8586
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 841 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-80
Query Match 76.0%; Score 19; DB 5; Length 841;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGTTCAGTGCAGCGAGAT 19
Db 275 GGTTCAGTGCAGCGAGAT 293
|||||
RESULT 38
PCT-US93-06251-81
; Sequence 81, Application PC/TUS9306251
; GENERAL INFORMATION:
; APPLICANT: Wickstrom, Eric and Rife, Jason P.
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06251
FILING DATE: 19930630
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8586
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 81:
SEQUENCE CHARACTERISTICS:
LENGTH: 841 base pairs

```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-81

Query Match      76.0%; Score 19; DB 5; Length 841;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
Db 275 GGTTCAGTGAGCCGAGAT 293

RESULT 39
US-09-257-179-11
; Sequence 11, Application US/09257179
; Patent No. 6410709
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 29 Human Secreted Proteins
; FILE REFERENCE: P2015P1
; CURRENT APPLICATION NUMBER: US/09/257,179
; CURRENT FILING DATE: 1999-02-25
; EARLIER APPLICATION NUMBER: PCT/US98/17709
; EARLIER FILING DATE: 1998-08-27
; EARLIER APPLICATION NUMBER: 60/056,270
; EARLIER FILING DATE: 1997-08-29
; EARLIER APPLICATION NUMBER: 60/056,271
; EARLIER FILING DATE: 1997-08-29
; EARLIER APPLICATION NUMBER: 60/056,247
; EARLIER FILING DATE: 1997-08-29
; EARLIER APPLICATION NUMBER: 60/056,073
; EARLIER FILING DATE: 1997-08-29
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 866
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (22)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (37)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-257-179-11

Query Match      76.0%; Score 19; DB 4; Length 866;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
Db 769 GGTTCAGTGAGCCGAGAT 787

RESULT 40
US-09-288-143-55
; Sequence 55, Application US/09288143
; Patent No. 6433139
; GENERAL INFORMATION:
; APPLICANT: Brewer et al.
; TITLE OF INVENTION: 53 Human Secreted Proteins
; FILE REFERENCE: P2018P1
; CURRENT APPLICATION NUMBER: US/09/288,143
; CURRENT FILING DATE: 1999-04-08
; EARLIER APPLICATION NUMBER: PCT/US98/21142
; EARLIER FILING DATE: 1998-10-08
; EARLIER APPLICATION NUMBER: 60/061,463

```

```

; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,529
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/071,498
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,527
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,536
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,532
; NUMBER OF SEQ ID NOS: 219
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 885
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (7)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-288-143-55

Query Match      76.0%; Score 19; DB 4; Length 885;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
Db 799 GGTTCAGTGAGCCGAGAT 817

RESULT 41
US-08-832-883-52/c
; Sequence 52, Application US/08832883
; Patent No. 5807681
; GENERAL INFORMATION:
; APPLICANT: Giordano, Antonio
; APPLICANT: Baldi, Alphonso
; TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND PROGNOSIS
; TITLE OF INVENTION: OF CANCER
; NUMBER OF SEQUENCES: 115
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
; STREET: Suite 1800 Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/832,883
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-13 US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 889 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-832-883-52

```

Query Match 76.0%; Score 19; DB 1; Length 889;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
|||||
Db 520 GGTTCAGTGAGCCGAGAT 502

RESULT 42

US-08-832-877-52/c
; Sequence 52, Application US/08832877
; Patent No. 5840506
; GENERAL INFORMATION:
; APPLICANT: Giordano, Antonio
; TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND PROGNOSIS OF
; TITLE OF INVENTION: CANCER
; NUMBER OF SEQUENCES: 116
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
; STREET: Suite 1800 Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/832,877
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-13 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 889 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-832-877-52

Query Match 76.0%; Score 19; DB 2; Length 889;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
|||||
Db 520 GGTTCAGTGAGCCGAGAT 502

RESULT 43

US-09-247-155-141
; Sequence 141, Application US/09247155A
; Patent No. 6312922
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Complementary DNAs
; FILE REFERENCE: GENSET.021A
; CURRENT APPLICATION NUMBER: US/09/247,155A
; CURRENT FILING DATE: 1999-02-09
; EARLIER APPLICATION NUMBER: 60/074,121

; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/081,563
; EARLIER FILING DATE: 1998-04-13
; EARLIER APPLICATION NUMBER: 60/096,116
; EARLIER FILING DATE: 1998-08-10
; EARLIER APPLICATION NUMBER: 60/099,273
; EARLIER FILING DATE: 1998-10-04
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: Patent.pm
; SEQ ID NO 141
; LENGTH: 891
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4..447
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 4..147
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 5.8999980926514
; OTHER INFORMATION: seq LILFFGKLLVVG/VG
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 858..863
; FEATURE:
; NAME/KEY: polyA_site
; LOCATION: 880..891
US-09-247-155-141

Query Match 76.0%; Score 19; DB 4; Length 891;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
|||||
Db 769 GGTTCAGTGAGCCGAGAT 787

RESULT 44

US-09-018-584A-40
; Sequence 40, Application US/09018584A
; Patent No. 6238863
; GENERAL INFORMATION:
; APPLICANT: Schumm, James W.
; APPLICANT: Bacher, Jeffery W.
; TITLE OF INVENTION: MATERIALS AND METHODS FOR
; TITLE OF INVENTION: IDENTIFYING AND ANALYZING
; NUMBER OF SEQUENCES: 147
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Promega Corporation
; STREET: 2800 Woods Hollow Road
; CITY: Madison
; STATE: Wisconsin
; COUNTRY: U.S.A.
; ZIP: 53711-5399
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Mb
; COMPUTER: IBM compatible PC
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Word 97 (DOS text format)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/018,584A
; FILING DATE: 04-Feb-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Grady J. Frenchick
; REGISTRATION NUMBER: 29,018
; REFERENCE/DOCKET NUMBER: 16026.9180
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 257-3501
; TELEFAX: (608) 257-2275

; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1000 bp
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Circular
; MOLECULE TYPE: Genomic DNA
; HYPOTHETICAL: no
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: S077
; US-09-018-584A-40

Query Match 76.0%; Score 19; DB 4; Length 1000;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTTGCAGTGAGCCGAGAT 19
|||||
Db 664 GGTTGCAGTGAGCCGAGAT 682

RESULT 45
US-09-641-638-636
; Sequence 636, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GENSET.051CE1
; CURRENT APPLICATION NUMBER: US/09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 636
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 10-508-191 : polymorphic base C or T
; NAME/KEY: misc_binding
; LOCATION: 481..500
; OTHER INFORMATION: 10-508-191.mis1, potential
; NAME/KEY: misc_binding
; LOCATION: 502..521
; OTHER INFORMATION: 10-508-191.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 311..328
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 727..746
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc_binding
; LOCATION: 489..513
; OTHER INFORMATION: 10-508-191 potential probe
US-09-641-638-636

Query Match 76.0%; Score 19; DB 4; Length 1000;
Best Local Similarity 100.0%; Pred. No. 0.012;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTTGCAGTGAGCCGAGAT 19
|||||
Db 140 GGTTGCAGTGAGCCGAGAT 158
Search completed: November 24, 2002, 15:03:42
Job time : 63.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2002, 14:30:54 ; Search time 52.5 Seconds
(without alignments)
180.348 Million cell updates/sec

Title: US-09-784-423-124
Perfect score: 25
Sequence: 1 GGTTCAGTGCAGCCGAGTAAGAGT 25

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 335578 seqs, 189365133 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671156

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	100.0	25	10	US-09-784-423-124
2	25	100.0	1000	10	US-09-784-423-32
3	20	80.0	384	10	US-09-867-701-2157
4	20	80.0	1875	12	US-10-052-586-389
5	20	80.0	2529	10	US-09-764-847-1859
6	20	80.0	5000	10	US-09-791-105-1
7	20	80.0	5131	10	US-09-764-877-3800
8	20	80.0	5139	10	US-09-764-877-3802
9	20	80.0	5814	10	US-09-764-847-1860
10	20	80.0	8886	10	US-09-764-878-412
11	20	80.0	8887	10	US-09-764-878-414
12	20	80.0	16106	10	US-09-764-877-2332
13	20	80.0	18878	10	US-09-764-877-3806
14	20	80.0	28588	10	US-09-764-887-399
15	20	80.0	31994	10	US-09-764-860-599
16	20	80.0	32248	10	US-09-764-864-1769
17	20	80.0	32248	10	US-09-764-864-1769
18	20	80.0	49984	10	US-09-739-457-5
19	20	80.0	65608	9	US-09-954-531-180

c 20	80.0	65608	10	US-09-962-436-292	Sequence 292, App
c 21	80.0	65608	10	US-09-962-832-119	Sequence 119, App
c 22	80.0	99014	10	US-09-880-107-3428	Sequence 3428, App
c 23	80.0	203654	10	US-09-820-905-3	Sequence 3, Appli
c 24	80.0	302250	10	US-09-962-832-154	Sequence 154, App
c 25	76.0	90	10	US-09-764-887-601	Sequence 601, App
c 26	76.0	95	10	US-09-764-869-2069	Sequence 2069, App
c 27	76.0	98	10	US-09-764-869-1767	Sequence 1767, App
c 28	76.0	98	10	US-09-764-869-2198	Sequence 2198, App
c 29	76.0	100	10	US-09-764-887-600	Sequence 600, App
c 30	76.0	102	10	US-09-764-877-3352	Sequence 3352, App
c 31	76.0	105	10	US-09-764-860-979	Sequence 979, App
c 32	76.0	107	10	US-09-764-869-2068	Sequence 2068, App
c 33	76.0	110	9	US-09-764-868-1442	Sequence 1442, App
c 34	76.0	114	10	US-09-764-860-792	Sequence 792, App
c 35	76.0	114	10	US-09-764-847-1408	Sequence 1408, App
c 36	76.0	114	10	US-09-764-847-1409	Sequence 1409, App
c 37	76.0	118	9	US-09-860-670-234	Sequence 234, App
c 38	76.0	120	10	US-09-764-860-659	Sequence 659, App
c 39	76.0	122	10	US-09-764-877-2168	Sequence 2168, App
c 40	76.0	123	10	US-09-764-847-1855	Sequence 1855, App
c 41	76.0	126	10	US-09-764-869-2373	Sequence 2373, App
c 42	76.0	129	10	US-09-783-590-10234	Sequence 10234, A
c 43	76.0	129	10	US-09-764-847-1975	Sequence 1975, App
c 44	76.0	130	10	US-09-764-877-4006	Sequence 4006, App
c 45	76.0	131	10	US-09-764-877-2688	Sequence 2688, App

ALIGNMENTS

RESULT 1
US-09-784-423-124
; Sequence 124, Application US/09784423
; Patent No. US20020012924A1
; GENERAL INFORMATION:
; APPLICANT: Schumm, James W.
; Bacher, Jeffery W.
; TITLE OF INVENTION: MATERIALS AND METHODS FOR
; IDENTIFYING AND ANALYZING INTERMEDIATE TANDEM
; REPEAT DNA MARKERS
; NUMBER OF SEQUENCES: 147
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Promega Corporation
; STREET: 2800 Woods Hollow Road
; CITY: Madison
; STATE: Wisconsin
; COUNTRY: U.S.A.
; ZIP: 53711-5399
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Mb
; COMPUTER: IBM compatible PC
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Word 97 (DOS text format)
; CURRENT APPLICATION NUMBER: US/09/784,423
; FILING DATE: 15-Feb-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/018,584
; FILING DATE: 04-Feb-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Grady J. Frenchick
; REGISTRATION NUMBER: 29,018
; REFERENCE/DOCKET NUMBER: 16026.9180
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 257-3501
; TELEFAX: (608) 257-2275
; INFORMATION FOR SEQ ID NO: 124
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25
; TYPE: Nucleic Acid
; STRANDEDNESS: Single


```

; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 124
US-09-784-423-124

Query Match      100.0%; Score 25; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 3.2e-06;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGGCCGAGATAAGAGT 25
    |||||||||||||||||||
Db 1 GGTTCAGTGGCCGAGATAAGAGT 25

RESULT 2
US-09-784-423-32
; Sequence 32, Application US/09784423
; Patent No. US20020012924A1
; GENERAL INFORMATION:
; APPLICANT: Schumm, James W.
; Bacher, Jeffery W.
; TITLE OF INVENTION: MATERIALS AND METHODS FOR
; IDENTIFYING AND ANALYZING INTERMEDIATE TANDEM
; REPEAT DNA MARKERS
; NUMBER OF SEQUENCES: 147
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Promega Corporation
; STREET: 2800 Woods Hollow Road
; CITY: Madison
; STATE: Wisconsin
; COUNTRY: U.S.A.
; ZIP: 53711-5399
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Mb
; COMPUTER: IBM compatible PC
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Word 97 (DOS text format)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/784.423
; FILING DATE: 15-Feb-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/018,584
; FILING DATE: 04-Feb-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Grady J. Frenchick
; REGISTRATION NUMBER: 29,018
; REFERENCE/DOCKET NUMBER: 16026.9180
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 257-3501
; TELEFAX: (608) 257-2275
; INFORMATION FOR SEQ ID NO: 32
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1000 bp
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Circular
; MOLECULE TYPE: Genomic DNA
; HYPOTHETICAL: no
; IMMEDIATE SOURCE:
; CLONE: S132
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 22
; SEQUENCE DESCRIPTION: SEQ ID NO: 32
US-09-784-423-32

Query Match      100.0%; Score 25; DB 10; Length 1000;
Best Local Similarity 100.0%; Pred. No. 3e-06;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGGCCGAGATAAGAGT 25
    |||||||||||||||||||
Db 441 GGTTCAGTGGCCGAGATAAGAGT 465

```

```

RESULT 3
US-09-867-701-2157/c
; Sequence 2157, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2157
; LENGTH: 384
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(384)
; OTHER INFORMATION: n = A,T,C or G
US-09-867-701-2157

Query Match      80.0%; Score 20; DB 10; Length 384;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGGCCGAGATA 20
    |||||||||||||||
Db 70 GGTTCAGTGGCCGAGATA 51

RESULT 4
US-10-052-586-389
; Sequence 389, Application US/10052586
; Patent No. US20020127584A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C1
; CURRENT APPLICATION NUMBER: US/10/052,586
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544

```

1	PRIOR FILING DATE: 1997-10-28
2	PRIOR APPLICATION NUMBER: 60/063564
3	PRIOR FILING DATE: 1997-10-28
4	PRIOR APPLICATION NUMBER: 60/063734
5	PRIOR FILING DATE: 1997-10-29
6	PRIOR APPLICATION NUMBER: 60/063870
7	PRIOR FILING DATE: 1997-10-31
8	PRIOR APPLICATION NUMBER: 60/064103
9	PRIOR FILING DATE: 1997-10-31
10	PRIOR APPLICATION NUMBER: 60/065311
11	PRIOR FILING DATE: 1997-11-13
12	PRIOR APPLICATION NUMBER: 60/066120
13	PRIOR FILING DATE: 1997-11-21
14	PRIOR APPLICATION NUMBER: 60/066466
15	PRIOR FILING DATE: 1997-11-24
16	PRIOR APPLICATION NUMBER: 60/066772
17	PRIOR FILING DATE: 1997-11-24
18	PRIOR APPLICATION NUMBER: 60/069335
19	PRIOR FILING DATE: 1997-12-11
20	PRIOR APPLICATION NUMBER: 60/069425
21	PRIOR FILING DATE: 1997-12-12
22	PRIOR APPLICATION NUMBER: 60/069870
23	PRIOR FILING DATE: 1997-12-17
24	PRIOR APPLICATION NUMBER: 60/068017
25	PRIOR FILING DATE: 1997-12-18
26	PRIOR APPLICATION NUMBER: 60/077450
27	PRIOR FILING DATE: 1998-03-10
28	PRIOR APPLICATION NUMBER: 60/077632
29	PRIOR FILING DATE: 1998-03-11
30	PRIOR APPLICATION NUMBER: 60/077649
31	PRIOR FILING DATE: 1998-03-11
32	PRIOR APPLICATION NUMBER: 60/078886
33	PRIOR FILING DATE: 1998-03-20
34	PRIOR APPLICATION NUMBER: 60/078939
35	PRIOR FILING DATE: 1998-03-20
36	PRIOR APPLICATION NUMBER: 60/079664
37	PRIOR FILING DATE: 1998-03-27
38	PRIOR APPLICATION NUMBER: 60/073786
39	PRIOR FILING DATE: 1998-03-27
40	PRIOR APPLICATION NUMBER: 60/080107
41	PRIOR FILING DATE: 1998-03-31
42	PRIOR APPLICATION NUMBER: 60/080194
43	PRIOR FILING DATE: 1998-03-31
44	PRIOR APPLICATION NUMBER: 60/080327
45	PRIOR FILING DATE: 1998-04-01
46	PRIOR APPLICATION NUMBER: 60/080333
47	PRIOR FILING DATE: 1998-04-01
48	PRIOR APPLICATION NUMBER: 60/081049
49	PRIOR FILING DATE: 1998-04-08
50	PRIOR APPLICATION NUMBER: 60/081070
51	PRIOR FILING DATE: 1998-04-08
52	PRIOR APPLICATION NUMBER: 60/081195
53	PRIOR FILING DATE: 1998-04-09
54	PRIOR APPLICATION NUMBER: 60/081838
55	PRIOR FILING DATE: 1998-04-15
56	PRIOR APPLICATION NUMBER: 60/082568
57	PRIOR FILING DATE: 1998-04-21
58	PRIOR APPLICATION NUMBER: 60/082569
59	PRIOR FILING DATE: 1998-04-21
60	PRIOR APPLICATION NUMBER: 60/082704
61	PRIOR FILING DATE: 1998-04-22
62	PRIOR APPLICATION NUMBER: 60/082797
63	PRIOR FILING DATE: 1998-04-22
64	PRIOR APPLICATION NUMBER: 60/083322
65	PRIOR FILING DATE: 1998-04-28
66	PRIOR APPLICATION NUMBER: 60/083495
67	PRIOR FILING DATE: 1998-04-29
68	PRIOR APPLICATION NUMBER: 60/083496
69	PRIOR FILING DATE: 1998-04-29
70	PRIOR APPLICATION NUMBER: 60/083499
71	PRIOR FILING DATE: 1998-04-29
72	PRIOR APPLICATION NUMBER: 60/083559
73	PRIOR FILING DATE: 1998-04-29

1	60/0843636	PRIOR APPLICATION NUMBER: 60/0843636
2	60/0843636	PRIOR FILING DATE: 1998-05-05
3	60/084414	PRIOR APPLICATION NUMBER: 60/084414
4	60/084414	PRIOR FILING DATE: 1998-05-06
5	60/084639	PRIOR APPLICATION NUMBER: 60/084639
6	60/084639	PRIOR FILING DATE: 1998-05-07
7	60/084640	PRIOR APPLICATION NUMBER: 60/084640
8	60/084640	PRIOR FILING DATE: 1998-05-07
9	60/084643	PRIOR APPLICATION NUMBER: 60/084643
10	60/084643	PRIOR FILING DATE: 1998-05-07
11	60/085573	PRIOR APPLICATION NUMBER: 60/085573
12	60/085573	PRIOR FILING DATE: 1998-05-15
13	60/085579	PRIOR APPLICATION NUMBER: 60/085579
14	60/085579	PRIOR FILING DATE: 1998-05-15
15	60/085580	PRIOR APPLICATION NUMBER: 60/085580
16	60/085580	PRIOR FILING DATE: 1998-05-15
17	60/085582	PRIOR APPLICATION NUMBER: 60/085582
18	60/085582	PRIOR FILING DATE: 1998-05-15
19	60/085700	PRIOR APPLICATION NUMBER: 60/085700
20	60/085700	PRIOR FILING DATE: 1998-05-15
21	60/086023	PRIOR APPLICATION NUMBER: 60/086023
22	60/086023	PRIOR FILING DATE: 1998-05-18
23	60/086392	PRIOR APPLICATION NUMBER: 60/086392
24	60/086392	PRIOR FILING DATE: 1998-05-22
25	60/086486	PRIOR APPLICATION NUMBER: 60/086486
26	60/086486	PRIOR FILING DATE: 1998-05-22
27	60/087098	PRIOR APPLICATION NUMBER: 60/087098
28	60/087098	PRIOR FILING DATE: 1998-05-28
29	60/087208	PRIOR APPLICATION NUMBER: 60/087208
30	60/087208	PRIOR FILING DATE: 1998-05-28
31	60/087609	PRIOR APPLICATION NUMBER: 60/087609
32	60/087609	PRIOR FILING DATE: 1998-06-02
33	60/087759	PRIOR APPLICATION NUMBER: 60/087759
34	60/087759	PRIOR FILING DATE: 1998-06-02
35	60/087827	PRIOR APPLICATION NUMBER: 60/087827
36	60/087827	PRIOR FILING DATE: 1998-06-03
37	60/088025	PRIOR APPLICATION NUMBER: 60/088025
38	60/088025	PRIOR FILING DATE: 1998-06-04
39	60/088028	PRIOR APPLICATION NUMBER: 60/088028
40	60/088028	PRIOR FILING DATE: 1998-06-04
41	60/088029	PRIOR APPLICATION NUMBER: 60/088029
42	60/088029	PRIOR FILING DATE: 1998-06-04
43	60/088033	PRIOR APPLICATION NUMBER: 60/088033
44	60/088033	PRIOR FILING DATE: 1998-06-04
45	60/088167	PRIOR APPLICATION NUMBER: 60/088167
46	60/088167	PRIOR FILING DATE: 1998-06-05
47	60/088202	PRIOR APPLICATION NUMBER: 60/088202
48	60/088202	PRIOR FILING DATE: 1998-06-05
49	60/088212	PRIOR APPLICATION NUMBER: 60/088212
50	60/088212	PRIOR FILING DATE: 1998-06-05
51	60/088217	PRIOR APPLICATION NUMBER: 60/088217
52	60/088217	PRIOR FILING DATE: 1998-06-05
53	60/088326	PRIOR APPLICATION NUMBER: 60/088326
54	60/088326	PRIOR FILING DATE: 1998-06-04
55	60/088655	PRIOR APPLICATION NUMBER: 60/088655
56	60/088655	PRIOR FILING DATE: 1998-06-09
57	60/088722	PRIOR APPLICATION NUMBER: 60/088722
58	60/088722	PRIOR FILING DATE: 1998-06-10
59	60/088738	PRIOR APPLICATION NUMBER: 60/088738
60	60/088738	PRIOR FILING DATE: 1998-06-10
61	60/088740	PRIOR APPLICATION NUMBER: 60/088740
62	60/088740	PRIOR FILING DATE: 1998-06-10
63	60/088811	PRIOR APPLICATION NUMBER: 60/088811
64	60/088811	PRIOR FILING DATE: 1998-06-10
65	60/088824	PRIOR APPLICATION NUMBER: 60/088824
66	60/088824	PRIOR FILING DATE: 1998-06-10
67	60/088825	PRIOR APPLICATION NUMBER: 60/088825
68	60/088825	PRIOR FILING DATE: 1998-06-10
69	60/088826	PRIOR APPLICATION NUMBER: 60/088826
70	60/088826	PRIOR FILING DATE: 1998-06-10
71	60/088861	PRIOR APPLICATION NUMBER: 60/088861
72	60/088861	PRIOR FILING DATE: 1998-06-11
73	60/088863	PRIOR APPLICATION NUMBER: 60/088863

```
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089090
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089908
```

```
Query Match      80.0%; Score 20; DB 12; Length 1875;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGTTCAGTGCAGCCGAGATA 20
      |||||
Db 1551 GGTTCAGTGCAGCCGAGATA 1570
```

RESULT 5

```
US-09-764-847-1859
; Sequence 1859, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764.847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1859
; LENGTH: 2529
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1859
```

```
Query Match      80.0%; Score 20; DB 10; Length 2529;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGTTCAGTGCAGCCGAGATA 20
      |||||
Db 850 GGTTCAGTGCAGCCGAGATA 869
```

RESULT 6

```
US-09-791-105-1
; Sequence 1, Application US/09791105
; Patent No. US200202225A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Epidauros Biotechnologie AG
; TITLE OF INVENTION: Method for detecting the presence of at least one single allele
; FILE REFERENCE: EPI 01/00
; CURRENT APPLICATION NUMBER: US/09/791.105
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: EP 00103844.7
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 5000
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-09-791-105-1
```

```
Query Match      80.0%; Score 20; DB 10; Length 5000;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGTTCAGTGCAGCCGAGATA 20
      |||||
Db 3838 GGTTCAGTGCAGCCGAGATA 3857
```

RESULT 7

```
US-09-764-877-3800/c
; Sequence 3800, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764.877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3800
; LENGTH: 5131
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3800
```

```
Query Match      80.0%; Score 20; DB 10; Length 5131;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGTTCAGTGCAGCCGAGATA 20
      |||||
Db 3049 GGTTCAGTGCAGCCGAGATA 3030
```

RESULT 8

```
US-09-764-877-3802/c
; Sequence 3802, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764.877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3802
; LENGTH: 5139
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3802
```

```
Query Match      80.0%; Score 20; DB 10; Length 5139;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGTTCAGTGCAGCCGAGATA 20
      |||||
Db 3049 GGTTCAGTGCAGCCGAGATA 3030
```

RESULT 9

```
US-09-764-847-1860
; Sequence 1860, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1860
; LENGTH: 5814
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1860

Query Match      80.0%; Score 20; DB 10; Length 5814;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGCAGTGAGCCGAGATA 20
   |||||||
Db 4138 GGTGCAGTGAGCCGAGATA 4157

RESULT 10
US-09-764-878-412
; Sequence 412, Application US/09764878
; Patent No. US20020090615A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAL21
; CURRENT APPLICATION NUMBER: US/09/764,878
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 428
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 412
; LENGTH: 8886
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-878-412

Query Match      80.0%; Score 20; DB 10; Length 8886;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGCAGTGAGCCGAGATA 20
   |||||||
Db 8094 GGTGCAGTGAGCCGAGATA 8113

RESULT 11
US-09-764-878-414
; Sequence 414, Application US/09764878
; Patent No. US20020090615A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAL21
; CURRENT APPLICATION NUMBER: US/09/764,878
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 428
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 414
; LENGTH: 8887
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-878-414

Query Match      80.0%; Score 20; DB 10; Length 8887;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GGTTCAGTGCAGCCGAGATA 20
   |||||||
Db 8094 GGTTCAGTGCAGCCGAGATA 8113

RESULT 12
US-09-764-877-2322/c
; Sequence 2322, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2322
; LENGTH: 16106
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2322

Query Match      80.0%; Score 20; DB 10; Length 16106;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGATA 20
   |||||||
Db 12684 GGTTCAGTGCAGCCGAGATA 12665

RESULT 13
US-09-764-877-3806/c
; Sequence 3806, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3806
; LENGTH: 18878
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3806

Query Match      80.0%; Score 20; DB 10; Length 18878;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGATA 20
   |||||||
Db 1960 GGTTCAGTGCAGCCGAGATA 1941

RESULT 14
US-09-764-887-399/c
; Sequence 399, Application US/09764887
; Patent No. US20020042096A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAL13
; CURRENT APPLICATION NUMBER: US/09/764,887
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
```

; NUMBER OF SEQ ID NOS: 658
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 399
; LENGTH: 28588
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-887-399

Query Match 80.0%; Score 20; DB 10; Length 28588;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGGCCGAGATA 20
|||||
Db 3891 GGTTCAGTGGCCGAGATA 3872

RESULT 15

US-09-764-860-599/c
; Sequence 599, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 599
; LENGTH: 31994
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-860-599

Query Match 80.0%; Score 20; DB 10; Length 31994;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGGCCGAGATA 20
|||||
Db 4422 GGTTCAGTGGCCGAGATA 4403

RESULT 16

US-09-764-864-1769
; Sequence 1769, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PP223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1769
; LENGTH: 32248
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-1769

Query Match 80.0%; Score 20; DB 10; Length 32248;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGGCCGAGATA 20
|||||
Db 12778 GGTTCAGTGGCCGAGATA 12797

RESULT 17

US-09-764-877-3487/c
; Sequence 3487, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3487
; LENGTH: 32248
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3487

Query Match 80.0%; Score 20; DB 10; Length 32248;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGGCCGAGATA 20
|||||
Db 19471 GGTTCAGTGGCCGAGATA 19452

RESULT 18

US-09-739-457-5/c
; Sequence 5, Application US/09739457
; Patent No. US20020072488A1
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady V. et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001014
; CURRENT APPLICATION NUMBER: US/09/739,457
; CURRENT FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 49984
; TYPE: DNA
; ORGANISM: Human
US-09-739-457-5

Query Match 80.0%; Score 20; DB 10; Length 49984;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTTCAGTGGCCGAGATA 20
|||||
Db 19071 GGTTCAGTGGCCGAGATA 19052

RESULT 19

US-09-954-531-180/c
; Sequence 180, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
; TITLE OF INVENTION: Gene Sets
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034

; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 180
; LENGTH: 65608
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(65608)
; OTHER INFORMATION: n=a,t,g or c
US-09-954-531-180

Query Match 80.0%; Score 20; DB 9; Length 65608;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATA 20
|||||
Db 62021 GGTTCAGTGAGCCGAGATA 62002

RESULT 20
US-09-962-436-292/c
; Sequence 292, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 292
; LENGTH: 65608
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: n=a,t,g or c
US-09-962-436-292

Query Match 80.0%; Score 20; DB 10; Length 65608;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATA 20
|||||
Db 62021 GGTTCAGTGAGCCGAGATA 62002

RESULT 21
US-09-962-832-119/c
; Sequence 119, Application US/09962832
; Patent No. US20020110821A1
; GENERAL INFORMATION:
; APPLICANT: Ederer, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-74
; CURRENT APPLICATION NUMBER: US/09/962,832
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,077

; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,280
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 119
; LENGTH: 65608
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: n=a,t,g or c
US-09-962-832-119

Query Match 80.0%; Score 20; DB 10; Length 65608;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATA 20
|||||
Db 62021 GGTTCAGTGAGCCGAGATA 62002

RESULT 22
US-09-880-107-3428/c
; Sequence 3428, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3428
; LENGTH: 99014
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U78027
US-09-880-107-3428

Query Match 80.0%; Score 20; DB 10; Length 99014;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATA 20
|||||
Db 93447 GGTTCAGTGAGCCGAGATA 93428

RESULT 23
US-09-820-905-3/c
; Sequence 3, Application US/09820905
; Patent No. US20020142938A1
; GENERAL INFORMATION:
; APPLICANT: Yan, Chunhua
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; FILE REFERENCE: CL001199
; CURRENT APPLICATION NUMBER: US/09/820,905
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0

```
; SEQ ID NO 3
; LENGTH: 203654
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(203654)
; OTHER INFORMATION: n = A,T,C or G
US-09-820-905-3

Query Match      80.0%; Score 20; DB 10; Length 203654;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATA 20
   |||||
Db 33496 GGTTCAGTGAGCCGAGATA 33477

RESULT 24
US-09-962-832-154/c
; Sequence 154, Application US/09962832
; Patent No. US20020110821A1
; GENERAL INFORMATION:
; APPLICANT: Eber, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 68290-74
; CURRENT APPLICATION NUMBER: US/09/962,832
; PRIOR APPLICATION NUMBER: US/60/235,077
; PRIOR FILING DATE: 2001-09-25
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,280
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 154
; LENGTH: 302250
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-832-154

Query Match      80.0%; Score 20; DB 10; Length 302250;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGATA 20
   |||||
Db 295292 GGTTCAGTGAGCCGAGATA 295273

RESULT 25
US-09-764-887-601/c
; Sequence 601, Application US/09764887
; Patent No. US20020042096A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P413
; CURRENT APPLICATION NUMBER: US/09/764,887
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 658
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 601
; LENGTH: 90
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-887-601

Query Match      76.0%; Score 19; DB 10; Length 90;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
   |||||
Db 75 GGTTCAGTGAGCCGAGAT 57

RESULT 26
US-09-764-869-2069
; Sequence 2069, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2069
; LENGTH: 95
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-2069

Query Match      76.0%; Score 19; DB 10; Length 95;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
   |||||
Db 20 GGTTCAGTGAGCCGAGAT 38

RESULT 27
US-09-764-869-1767/c
; Sequence 1767, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1767
; LENGTH: 98
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-1767

Query Match      76.0%; Score 19; DB 10; Length 98;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
   |||||
Db 75 GGTTCAGTGAGCCGAGAT 57

RESULT 28
US-09-764-869-2198
; Sequence 2198, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
```

; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2198
; LENGTH: 98
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-2198

Query Match 76.0%; Score 19; DB 10; Length 98;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGAT 19
|||||
Db 24 GGTTCAGTGCAGCCGAGAT 42

RESULT 29
US-09-764-887-600/c
; Sequence 600, Application US/09764887
; Patent No. US20020042096A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA113
; CURRENT APPLICATION NUMBER: US/09/764,887
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 658
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 600
; LENGTH: 100
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-887-600

Query Match 76.0%; Score 19; DB 10; Length 100;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGAT 19
|||||
Db 80 GGTTCAGTGCAGCCGAGAT 62

RESULT 30
US-09-764-877-3352/c
; Sequence 3352, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3352
; LENGTH: 102
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3352

Query Match 76.0%; Score 19; DB 10; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGAT 19
|||||
Db 40 GGTTCAGTGCAGCCGAGAT 22

RESULT 31
US-09-764-860-979
; Sequence 979, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 979
; LENGTH: 105
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-860-979

Query Match 76.0%; Score 19; DB 10; Length 105;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTTCAGTGCAGCCGAGATA 20
|||||
Db 43 GTTCAGTGCAGCCGAGATA 61

RESULT 32
US-09-764-869-2068
; Sequence 2068, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2068
; LENGTH: 107
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-2068

Query Match 76.0%; Score 19; DB 10; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGAT 19
|||||
Db 23 GGTTCAGTGCAGCCGAGAT 41

RESULT 33
US-09-764-868-1442/c
; Sequence 1442, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT332
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1442
; LENGTH: 110
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-764-868-1442

Query Match 76.0%; Score 19; DB 9; Length 110;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
|||||
DB 83 GGTTCAGTGAGCCGAGAT 55

RESULT 34

US-09-764-860-792/c
; Sequence 792, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 792
; LENGTH: 114
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-860-792

Query Match 76.0%; Score 19; DB 10; Length 114;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
|||||
DB 73 GGTTCAGTGAGCCGAGAT 55

RESULT 35

US-09-764-847-1408/c
; Sequence 1408, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1408
; LENGTH: 114
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1408

Query Match 76.0%; Score 19; DB 10; Length 114;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
|||||
DB 73 GGTTCAGTGAGCCGAGAT 55

RESULT 36

US-09-764-847-1409/c
; Sequence 1409, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1409
; LENGTH: 114
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1409

Query Match 76.0%; Score 19; DB 10; Length 114;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
|||||
DB 73 GGTTCAGTGAGCCGAGAT 55

RESULT 37

US-09-860-670-234
; Sequence 234, Application US/09860670
; Patent No. US20020165137A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA127P1
; CURRENT APPLICATION NUMBER: US/09/860,670
; CURRENT FILING DATE: 2001-05-21
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 234
; LENGTH: 118
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-860-670-234

Query Match 76.0%; Score 19; DB 9; Length 118;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGAGCCGAGAT 19
|||||
DB 50 GGTTCAGTGAGCCGAGAT 58

RESULT 38

US-09-764-860-659/c
; Sequence 659, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 659
; LENGTH: 120
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-860-659

Query Match 76.0%; Score 19; DB 10; Length 120;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 14 GGTTCAGTGCAGCCGAGAT 32

RESULT 43

US-09-764-847-1975/c
; Sequence 1975, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1975
; LENGTH: 129
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1975

Query Match 76.0%; Score 19; DB 10; Length 129;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGAT 19

Db 81 GGTTCAGTGCAGCCGAGAT 63

RESULT 44

US-09-764-877-4006
; Sequence 4006, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4006
; LENGTH: 130
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-4006

Query Match 76.0%; Score 19; DB 10; Length 130;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGAT 19

Db 70 GGTTCAGTGCAGCCGAGAT 88

RESULT 45

US-09-764-877-2688/c
; Sequence 2688, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2688

; LENGTH: 131
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2688

Query Match 76.0%; Score 19; DB 10; Length 131;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTTCAGTGCAGCCGAGAT 19

Db 81 GGTTCAGTGCAGCCGAGAT 63

Search completed: November 24, 2002, 16:37:43
Job time : 203.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2002, 13:30:43 ; Search time 46.5 Seconds
(without alignments)
164.880 Million cell updates/sec

Title: US-09-784-423-125

Perfect score: 25

Sequence: 1 TGTGCCAGGACCAAGAAATTACAG 25

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	25	100.0	25	4	US-09-018-584A-125	Sequence 125, App
c 2	25	100.0	1000	4	US-09-018-584A-32	Sequence 32, Appl
3	15	60.0	28720	4	US-09-341-587-7	Sequence 7, Appl
4	14	56.0	627	4	US-08-981-030-2	Sequence 2, Appl
5	14	56.0	627	4	US-08-981-030-12	Sequence 12, Appl
6	14	56.0	682	4	US-08-981-030-1	Sequence 1, Appl
7	14	56.0	1727	1	US-08-289-458-3	Sequence 3, Appl
8	14	56.0	1727	2	US-08-761-549-3	Sequence 3, Appl
9	14	56.0	1727	4	US-09-127-646-3	Sequence 3, Appl
10	14	56.0	4376	1	US-08-119-125A-1	Sequence 1, Appl
11	14	56.0	6744	1	US-08-119-125A-2	Sequence 2, Appl
12	14	56.0	9493	2	US-08-639-857-23	Sequence 23, Appl
c 14	14	56.0	9493	4	US-08-469-260A-163	Sequence 163, App
c 15	13	52.0	22	2	US-09-798-096-10	Sequence 10, Appl
c 16	13	52.0	56	1	US-08-332-766A-76	Sequence 76, Appl
c 17	13	52.0	56	1	US-08-229-279-6	Sequence 6, Appl
c 18	13	52.0	56	1	US-08-701-269-6	Sequence 6, Appl
c 19	13	52.0	89	4	US-09-511-625B-57	Sequence 57, Appl
c 20	13	52.0	91	4	US-09-084-120-23	Sequence 23, Appl
c 21	13	52.0	445	2	US-08-332-766A-16	Sequence 16, Appl
c 22	13	52.0	611	4	US-09-328-111-416	Sequence 416, Appl
c 23	13	52.0	646	4	US-08-998-416-77	Sequence 77, Appl
c 24	13	52.0	691	2	US-08-365-486A-29	Sequence 29, Appl
c 25	13	52.0	691	4	US-08-880-342-29	Sequence 29, Appl
26	13	52.0	700	4	US-08-991-789A-174	Sequence 174, App
27	13	52.0	700	4	US-09-062-451-174	Sequence 174, App
					US-09-598-326-174	Sequence 174, App

28 13 52.0 829 3 US-08-961-083-133 Sequence 133, App
c 29 13 52.0 836 1 US-08-554-612C-26 Sequence 26, Appl
c 30 13 52.0 932 1 US-08-554-612C-20 Sequence 20, Appl
c 31 13 52.0 1001 4 US-09-641-638-514 Sequence 514, App
c 32 13 52.0 1001 4 US-09-641-638-515 Sequence 515, App
c 33 13 52.0 1001 4 US-09-641-638-516 Sequence 516, App
c 34 13 52.0 1007 1 US-08-554-612C-17 Sequence 17, Appl
c 35 13 52.0 1007 1 US-08-554-612C-18 Sequence 18, Appl
c 36 13 52.0 1356 4 US-09-276-531-129 Sequence 129, App
c 37 13 52.0 1508 1 US-08-554-612C-16 Sequence 16, Appl
c 38 13 52.0 1701 3 US-08-599-968-2 Sequence 2, Appl
c 39 13 52.0 1743 4 US-08-487-596-9 Sequence 9, Appl
c 40 13 52.0 1758 5 PCT-US92-01015-1 Sequence 1, Appl
c 41 13 52.0 1782 4 US-09-149-476-120 Sequence 120, App
c 42 13 52.0 1981 3 US-09-017-706-3 Sequence 3, Appl
c 43 13 52.0 1981 3 US-09-017-706-4 Sequence 4, Appl
c 44 13 52.0 1981 3 US-09-017-706-5 Sequence 5, Appl
c 45 13 52.0 1981 3 US-09-017-706-6 Sequence 6, Appl

ALIGNMENTS

RESULT 1

US-09-018-584A-125
; Sequence 125, Application US/09018584A
; Patent No. 6238863

GENERAL INFORMATION:

APPLICANT: Schumm, James W.

APPLICANT: Bacher, Jeffery W.

TITLE OF INVENTION: MATERIALS AND METHODS FOR

TITLE OF INVENTION: IDENTIFYING AND ANALYZING INTERMEDIATE TANDEM

TITLE OF INVENTION: REPEAT DNA MARKERS

NUMBER OF SEQUENCES: 147

CORRESPONDENCE ADDRESS:

ADDRESSEE: Promega Corporation

STREET: 2800 Woods Hollow Road

CITY: Madison

STATE: Wisconsin

COUNTRY: U.S.A.

ZIP: 53711-5399

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Mb

COMPUTER: IBM compatible PC

OPERATING SYSTEM: Windows 95

SOFTWARE: Word 97 (DOS text format)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/018,584A

FILING DATE: 04-Feb-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Grady J. Frenchick

REGISTRATION NUMBER: 29,018

REFERENCE/DOCKET NUMBER: 16026.9180

TELECOMMUNICATION INFORMATION:

TELEPHONE: (608) 257-3501

TELEFAX: (608) 257-2275

INFORMATION FOR SEQ ID NO: 125:

SEQUENCE CHARACTERISTICS:

LENGTH: 25

TYPE: Nucleic Acid

STRANDEDNESS: Single

TOPOLOGY: Linear

US-09-018-584A-125

Query Match 100.0%; Score 25; DB 4; Length 25;

Best Local Similarity 100.0%; Pred. No. 9.3e-06;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGACCAAGAAATTACAG 25

|||||

Db 1 TGTGCCAGGACCAAGAAATTACAG 25

```
RESULT 2
US-09-018-584A-32/C
; Sequence 32, Application US/09018584A
; Patent No. 6238863
; GENERAL INFORMATION:
; APPLICANT: Schumm, James W.
; APPLICANT: Bacher, Jeffrey W.
; TITLE OF INVENTION: MATERIALS AND METHODS FOR
; TITLE OF INVENTION: IDENTIFYING AND ANALYZING INTERMEDIATE TANDEM
; TITLE OF INVENTION: REPEAT DNA MARKERS
; NUMBER OF SEQUENCES: 147
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Promega Corporation
; STREET: 2800 Woods Hollow Road
; CITY: Madison
; STATE: Wisconsin
; COUNTRY: U.S.A.
; ZIP: 53711-5399
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 MB
; COMPUTER: IBM compatible PC
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Word 97 (DOS text format)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/018,584A
; FILING DATE: 04-Feb-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Grady J. Frenchick
; REGISTRATION NUMBER: 29,018
; REFERENCE/DOCKET NUMBER: 16026.9180
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 257-3501
; TELEFAX: (608) 257-2275
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1000 bp
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Circular
; MOLECULE TYPE: Genomic DNA
; HYPOTHETICAL: NO
; IMMEDIATE SOURCE:
; CLONE: S132
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 22
US-09-018-584A-32
Query Match 100.0%; Score 25; DB 4; Length 1000;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGTGCCAGGACCAAGAAATTACAG 25
Db 726 TGTGCCAGGACCAAGAAATTACAG 702
;
RESULT 3
US-09-341-587-7
; Sequence 7, Application US/09341587
; Patent No. 6346606
; GENERAL INFORMATION:
; APPLICANT: Mollenhauer, Jan
; TITLE OF INVENTION: Protein Containing an SRCR Domain
; FILE REFERENCE: 4121-108
; CURRENT APPLICATION NUMBER: US/09/341,587
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: PCT/DE98/00096
; EARLIER FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
```

```
; LENGTH: 28720
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-341-587-7
Query Match 60.0%; Score 15; DB 4; Length 28720;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 CAGGAACACAGAAATT 20
Db 5047 CAGGAACACAGAAATT 5061
;
RESULT 4
US-08-981-030-2
; Sequence 2, Application US/08981030
; Patent No. 6447783
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: FGF9 AS A SPECIFIC LIGAND FOR FGFR3
; NUMBER OF SEQUENCES: 13
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/981,030
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/IL96/00011
; FILING DATE: 12-JUN-1996
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 627 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Mus pahari
US-08-981-030-2
Query Match 56.0%; Score 14; DB 4; Length 627;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 ACCAGAAATTAC 24
Db 542 ACCAGAAATTAC 555
;
RESULT 5
US-08-981-030-12
; Sequence 12, Application US/08981030
; Patent No. 6447783
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: FGF9 AS A SPECIFIC LIGAND FOR FGFR3
; NUMBER OF SEQUENCES: 13
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/981,030
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/IL96/00011
```



```

; TITLE OF INVENTION: PLANT Gp2 PROMOTERS AND USES THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Stewart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/761,549
; FILING DATE: 06-DEC-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/289,458
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 12176-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1727 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CAAT_signal
; LOCATION: 1100..1103
; FEATURE:
; NAME/KEY: TATA_signal
; LOCATION: 1139..1146
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1169
; OTHER INFORMATION: /note= "Transcriptional start site"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1172
; OTHER INFORMATION: /note= "pgp50 5' end"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1234..1236
; OTHER INFORMATION: /note= "Translation start codon"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1298
; OTHER INFORMATION: /note= "Intron start site"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1233
; OTHER INFORMATION: /note= "Gp2 promoter sequence"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..1727
; OTHER INFORMATION: /note= "Gp2 Genomic DNA clone"
; US-08-761-549-3
;
; Query Match 56.0%; Score 14; DB 2; Length 1727;
; Best Local Similarity 100.0%; Pred. No. 19;
; Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 8 GGAACCAAGAAATT 21
; |||||

```

```

Db 1619 GGAACCAAGAAATT 1632
;
; RESULT 9
; US-09-127-646-3
; Sequence 3, Application US/09127646
; Patent No. 6291744
; GENERAL INFORMATION:
; APPLICANT: Baden, Catherine S.
; APPLICANT: Dunsmuir, Pamela
; APPLICANT: Lee, Kathleen Y.
; APPLICANT: DNA Plant Technology Corporation
; TITLE OF INVENTION: Nucleic Acids Encoding Plant Group 2 Proteins and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 012176-004020US
; CURRENT APPLICATION NUMBER: US/09/127,646
; CURRENT FILING DATE: 1998-07-31
; EARLIER APPLICATION NUMBER: US 08/289,458
; EARLIER FILING DATE: 1994-08-12
; EARLIER APPLICATION NUMBER: US 08/761,549
; EARLIER FILING DATE: 1996-12-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1727
; TYPE: DNA
; ORGANISM: Capsicum annuum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1727)
; OTHER INFORMATION: pepper plant Group 2 (Gp2) genomic DNA clone
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1233)
; OTHER INFORMATION: Gp2 promoter sequence
; FEATURE:
; NAME/KEY: CAAT_signal
; LOCATION: (1100)..(1103)
; FEATURE:
; NAME/KEY: TATA_signal
; LOCATION: (1139)..(1146)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1169)
; OTHER INFORMATION: transcriptional start site
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1172)
; OTHER INFORMATION: pgp50 5' end
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1234)..(1236)
; OTHER INFORMATION: translation start codon
; FEATURE:
; NAME/KEY: intron
; LOCATION: (1298)
; OTHER INFORMATION: Intron start site
; US-09-127-646-3
;
; Query Match 56.0%; Score 14; DB 4; Length 1727;
; Best Local Similarity 100.0%; Pred. No. 19;
; Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 8 GGAACCAAGAAATT 21
; |||||
; Db 1619 GGAACCAAGAAATT 1632
;
; RESULT 10
; US-08-119-125A-1
; Sequence 1, Application US/08119125A
; Patent No. 5610011
; GENERAL INFORMATION:

```

APPLICANT: SMITH, Hilda Elizabeth
 APPLICANT: VECHT, Uri
 TITLE OF INVENTION: DNA sequences which code for virulence
 TITLE OF INVENTION: Characteristics of Streptococcus suis and parts thereof, polypeptides
 TITLE OF INVENTION: antibodies derived therefrom and the use thereof for the diagnosis
 TITLE OF INVENTION: protection against infection by S. suis in mammals, including
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Centraal Diergeneeskundig Instituut
 STREET: Edelhertweg 15
 CITY: PH Lelystad
 STATE:
 COUNTRY: The Netherlands
 ZIP: NL-8219
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
 COMPUTER: IBM compatible
 OPERATING SYSTEM: MS-DOS v.6.0
 SOFTWARE: WordPerfect v. 6.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/119,125A
 FILING DATE: 20-SEP-1993
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/NL92/00054
 FILING DATE: 19-MAR-1992
 APPLICATION NUMBER: NL 9100510
 FILING DATE: 21-MAR-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Handal, Anthony H.
 REGISTRATION NUMBER: 26275
 REFERENCE/DOCKET NUMBER: SMITHHE119125
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (203) 838-8589
 TELEFAX: (203) 838-8794
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4376 base pairs
 TYPE: Nucleic acid with corresponding amino acids
 STRANDEDNESS: single stranded
 TOPOLOGY: linear
 MOLECULE TYPE: genomic DNA
 ORIGINAL SOURCE:
 ORGANISM: Streptococcus suis type II (pathogenic)
 FEATURE:
 OTHER INFORMATION: Extracellular protein factor (EF) gene
 FEATURE:
 NAME/KEY: promoter -35 region
 LOCATION: bp 66 to 71
 FEATURE:
 NAME/KEY: promoter -10 region
 LOCATION: bp 89 to 94
 FEATURE:
 NAME/KEY: promoter -35 region
 LOCATION: bp 153 to 158
 FEATURE:
 NAME/KEY: promoter -10 region
 LOCATION: bp 176 to 181
 FEATURE:
 NAME/KEY: ribosome binding site
 LOCATION: bp 350 to 356
 FEATURE:
 NAME/KEY: signal peptide
 LOCATION: bp 361 to 498
 FEATURE:
 NAME/KEY: mature peptide
 LOCATION: bp 499 to 2890
 FEATURE:
 NAME/KEY: dyad symmetry regions
 LOCATION: from bp 4186 to 4198 and from bp 4203 to 4215
 NAME/KEY: dyad symmetry regions
 LOCATION: from bp 4243 to 4257 and from bp 4263 to 4276

```

US-08-119-125A-1
Query Match          56.0%; Score 14; DB 1; Length 4376;
Best Local Similarity 100.0%; Pred. NO. 19;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6 CAGGAACCAAGAAAT 19
        |||||
Db      314 CAGGAACCAAGAAAT 327

RESULT 11
US-08-119-125A-2
; Sequence 2, Application US/08119125A
; Patent No. 5610011
; GENERAL INFORMATION:
; APPLICANT: SMITH, Hilda Elizabeth
; APPLICANT: VECHT, Uri
; TITLE OF INVENTION: DNA Sequences which code for Virulence
; TITLE OF INVENTION: Characteristics of Streptococcus suis and parts thereof, po
; TITLE OF INVENTION: antibodies derived therefrom and the use thereof for the di
; TITLE OF INVENTION: protection against infection by S. suis in mammals, includi
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Centraal Diergeneeskundig Instituut
; STREET: Edelhertweg 15
; CITY: PH Lelystad
; STATE:
; COUNTRY: The Netherlands
; ZIP: NL-8219
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS v.6.0
; SOFTWARE: WordPerfect v. 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/119,125A
; FILING DATE: 20-SEP-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/NL92/00054
; FILING DATE: 19-MAR-1992
; APPLICATION NUMBER: NL 9100510
; FILING DATE: 21-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Handal, Anthony H.
; REGISTRATION NUMBER: 26275
; REFERENCE/DOCKET NUMBER: SMITHHE119125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203) 838-8589
; TELEFAX: (203) 838-8794
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6744 base pairs
; TYPE: Nucleic acid with corresponding amino acids
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus suis type II (pathogenic)
; FEATURE:
; OTHER INFORMATION: Extracellular factor related protein (EF*) gene
; FEATURE:
; NAME/KEY: promoter -35 region
; LOCATION: bp 66 to 71
; FEATURE:
; NAME/KEY: promoter -10 region
; LOCATION: bp 89 to 94
; FEATURE:
; NAME/KEY: promoter -35 region
; LOCATION: bp 153 to 158
; FEATURE:
; NAME/KEY: promoter -10 region

```



```
/ LOCATION: bp 176 to 181
/ FEATURE:
/ NAME/KEY: ribosome binding site
/ LOCATION: bp 350 to 356
/ FEATURE:
/ NAME/KEY: signal peptide
/ LOCATION: bp 361 to 498
/ FEATURE:
/ NAME/KEY: start of repetitive units R1-R11
/ LOCATION: bp 4869, 3097, 3292, 3520, 4087, 4381, 4609, 4837,
/ LOCATION: 5065, 5293, 5521:
/ FEATURE:
/ NAME/KEY: start of repetitive Asn-Pro-Asn-Leu sequences
/ LOCATION: bp 2932, 3160, 3355, 3583, 4150, 4444, 4672, 4900,
/ LOCATION: 5128, 5356, 5584:
/ FEATURE:
/ NAME/KEY: dyad symmetry regions
/ LOCATION: from bp 6554 to 6566 and from bp 6571 to 6583
/ FEATURE:
/ NAME/KEY: dyad symmetry regions
/ LOCATION: from bp 6611 to 6625 and from bp 6631 to 6644
US-08-119-125A-2
Query Match 56.0%; Score 14; DB 1; Length 6744;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CAGGACCCAGAAAT 19
|||||
DB 314 CAGGACCCAGAAAT 327

RESULT 12
US-08-639-857-23
; Sequence 23, Application US/08639857
; Patent No. 595318
; GENERAL INFORMATION:
; APPLICANT: Simons, J. N.
; APPLICANT: Desai, S. M.
; APPLICANT: Mushahwar I. K.
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL FOR CONTROLLING THE
; TITLE OF INVENTION: TRANSLATION OF HEPATITIS GB PROTEINS
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Rd
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/639,857
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Foremski, Priscilla E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5793.US.P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-0378
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9493 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
```

```
US-08-639-857-23
Query Match 56.0%; Score 14; DB 2; Length 9493;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGAACCA 14
|||||
DB 4037 TGTGCCAGGAACCA 4050

RESULT 13
US-08-469-260A-163
; Sequence 163, Application US/08469260A
; Patent No. 6451578
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUERHOFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUIJK
; APPLICANT: ISA K. MUSHAHWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,260A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,550
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5527.PC.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 163:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9493 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-469-260A-163

Query Match 56.0%; Score 14; DB 4; Length 9493;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGAACCA 14
|||||
DB 4037 TGTGCCAGGAACCA 4050
```

RESULT 14
US-09-798-096-10/c
; Sequence 10, Application US/09798096
; Patent No. 6398378
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF RECOL2 EXPRESSION
; FILE REFERENCE: RTS-0207
; CURRENT APPLICATION NUMBER: US/09/798,096
; CURRENT FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 10
; LENGTH: 99500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-798-096-10

Query Match 56.0%; Score 14; DB 4; Length 99500;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGAACCA 14
|||||
DB 31468 TGTGCCAGGAACCA 31455

RESULT 15
US-08-332-766A-76/c
; Sequence 76, Application US/08332766A
; Patent No. 5843647
; GENERAL INFORMATION:
; APPLICANT: JEFFREYS, Alec J.
; APPLICANT: ARMOUR, John
; TITLE OF INVENTION: SIMPLE TANDEM REPEATS
; NUMBER OF SEQUENCES: 125
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABRY & CUSHMAN, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332,766A
; FILING DATE: 01-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9326052.9
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BIRD, Donald J.
; REGISTRATION NUMBER: 25,323
; REFERENCE/DOCKET NUMBER: 217211/M94/0434/GB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 76:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-332-766A-76

Query Match 52.0%; Score 13; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 ACCAGAAATTTC 23
|||||
DB 21 ACCAGAAATTTC 9
|||||
RESULT 16
US-08-229-279-6/c
; Sequence 6, Application US/08229279
; Patent No. 5648211
; GENERAL INFORMATION:
; APPLICANT: Fraiser, Melinda S.
; APPLICANT: Spargo, Catherine A.
; APPLICANT: Walker, George T.
; APPLICANT: Van Cleve, Mark
; APPLICANT: Wright, David J.
; TITLE OF INVENTION: STRAND DISPLACEMENT AMPLIFICATION USING
; TITLE OF INVENTION: THERMOPHILIC ENZYMES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
; COMPANY: Company
; STREET: 1 Becton Drive
; CITY: Franklin Lakes
; STATE: NJ
; COUNTRY: US
; ZIP: 07417
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/229,279
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fugit, Donna R.
; REGISTRATION NUMBER: 32,135
; REFERENCE/DOCKET NUMBER: P-2961
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 56 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-229-279-6

Query Match 52.0%; Score 13; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 CCAGGAACAGAA 17
|||||
DB 19 CCAGGAACAGAA 7
|||||

RESULT 17
US-08-701-269-6/c
; Sequence 6, Application US/08701269
; Patent No. 5744311
; GENERAL INFORMATION:
; APPLICANT: Fraiser, Melinda S.
; APPLICANT: Spargo, Catherine A.
; APPLICANT: Walker, George T.
; APPLICANT: Van Cleve, Mark
; APPLICANT: Wright, David J.
; TITLE OF INVENTION: STRAND DISPLACEMENT AMPLIFICATION USING
; TITLE OF INVENTION: THERMOPHILIC ENZYMES

```
/ NUMBER OF SEQUENCES: 19
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
/ ADDRESSEE: Company
/ STREET: 1 Becton Drive
/ CITY: Franklin Lakes
/ STATE: NJ
/ COUNTRY: US
/ ZIP: 07417
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/701.269
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/229,279
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fugit, Donna R.
/ REGISTRATION NUMBER: 32,135
/ REFERENCE/DOCKET NUMBER: P-2961
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 56 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
US-08-701-269-6

Query Match          52.0%; Score 13; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CCAGGACCAGAA 17
   |||||
DB 19 CCAGGACCAGAA 7

RESULT 18
US-09-511-625B-57/c
; Sequence 57, Application US/09511625B
; Patent No. 6368828
; GENERAL INFORMATION:
; APPLICANT: LaRoche, William J.
; APPLICANT: Patel, Bharvin
; APPLICANT: Pierce, Jacalyn H.
; TITLE OF INVENTION: ATTENUATED AND DOMINANT NEGATIVE VARIANT
; TITLE OF INVENTION: CDNAS OF STAT6: STAT6B AND STAT6C
; FILE REFERENCE: 14014.0300u1
; CURRENT APPLICATION NUMBER: US/09/511,625B
; CURRENT FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: PCT/US98/17821
; PRIOR FILING DATE: 1998-08-27
; PRIOR APPLICATION NUMBER: 60/070,397
; PRIOR FILING DATE: 1998-01-05
; PRIOR APPLICATION NUMBER: 60/056,075
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 89
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of artificial Sequence:/No. 6368828e -
; OTHER INFORMATION: synthetic construct
US-09-511-625B-57
```

```
Query Match          52.0%; Score 13; DB 4; Length 89;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TCCAGGACCAG 15
   |||||
DB 88 TCCAGGACCAG 76

RESULT 19
US-09-084-120-23/c
; Sequence 23, Application US/09084120
; Patent No. 6251592
; GENERAL INFORMATION:
; APPLICANT: TANG, Jianqing
; APPLICANT: MELANCON, Serge B.
; TITLE OF INVENTION: A NOVEL STR MARKER SYSTEM
; TITLE OF INVENTION: FOR DNA FINGERPRINTING
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SWABEY OGILVY RENAULT
; STREET: Suite 1600, 1981 McGill College Avenue
; CITY: Montreal
; STATE: QC
; COUNTRY: Canada
; ZIP: H3A 2Y3
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/084,120
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: COTE, France
; REGISTRATION NUMBER: 37,037
; REFERENCE/DOCKET NUMBER: 13251-40US FC/1d
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 514 845-7126
; TELEFAX: 514 288-8389
; TELEX:
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 91 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-09-084-120-23

Query Match          52.0%; Score 13; DB 4; Length 91;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GCCAGGACCAGA 16
   |||||
DB 51 GCCAGGACCAGA 39

RESULT 20
US-08-332-766A-16/c
; Sequence 16, Application US/08332766A
; Patent No. 5843647
; GENERAL INFORMATION:
; APPLICANT: JEFFREYS, Alec J.
; APPLICANT: ARMOUR, John
; TITLE OF INVENTION: SIMPLE TANDEM REPEATS
; NUMBER OF SEQUENCES: 125
```

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: CUSHMAN DABBY & CUSHMAN, L.I.P.
;; STREET: 1100 New York Avenue, N.W.
;; CITY: Washington
;; STATE: D. C.
;; COUNTRY: U.S.A.
;; ZIP: 20005-3918
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/332,766A
;; FILING DATE: 01-NOV-1994
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: GB 9326052.9
;; FILING DATE: 21-DEC-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BIRD, Donald J.
;; REGISTRATION NUMBER: 25,323
;; REFERENCE/DOCKET NUMBER: 217211/M94/0434/GB
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 861-3000
;; TELEFAX: (202) 822-0944
;; TELEX: 6714627 CUSH
;; INFORMATION FOR SEQ ID NO: 16:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 445 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; US-08-332-766A-16

Query Match 52.0%; Score 13; DB 2; Length 445;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ACCAGAAATTAC 23
|||||
DB 235 ACCAGAAATTAC 223

RESULT 21
US-09-328-111-416/c
; Sequence 416, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 416
; LENGTH: 611

;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(611)
;; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-416

Query Match 52.0%; Score 13; DB 4; Length 611;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTCCAGGAACCA 14
|||||
DB 267 GTCCAGGAACCA 255

RESULT 22
US-08-998-416-77
; Sequence 77, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reibschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 77:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 646 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1023UP
; US-08-998-416-77

Query Match 52.0%; Score 13; DB 4; Length 646;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AGGAACCAAGAAAT 19
| | | | | | | | | |
Db 81 AGGAACCAAGAAAT 93

RESULT 23

US-08-365-486A-29/c
; Sequence 29, Application US/08365486A
; Patent No. 5834306
; GENERAL INFORMATION:

APPLICANT: Webster, Keith A.
APPLICANT: Bishopric, Nanette H.
TITLE OF INVENTION: Tissue Specific Hypoxia Regulated
TITLE OF INVENTION: Therapeutic Constructs
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/365,486A
FILING DATE: 23-DEC-1994
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 8255-0018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 691 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO

ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: human enolase gene (EMBL #X56832)

INDIVIDUAL ISOLATE: fragment containing nucleotides -628 to +63
FEATURE:
NAME/KEY: CDS
LOCATION: 629..691

US-08-365-486A-29

Query Match 52.0%; Score 13; DB 2; Length 691;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AGGAACCAAGAAAT 19
| | | | | | | | | |
Db 581 AGGAACCAAGAAAT 569

RESULT 24

US-08-880-342-29/c
; Sequence 29, Application US/08880342
; Patent No. 6218179
; GENERAL INFORMATION:

APPLICANT: Webster, Keith A.
APPLICANT: Bishopric, Nanette H.
APPLICANT: Murphy, Brian
APPLICANT: Laderoute, Keith R.

APPLICANT: Green, Christopher J.
TITLE OF INVENTION: Tissue Specific Hypoxia Regulated
TITLE OF INVENTION: Therapeutic Constructs
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:

ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,342
FILING DATE: 23-JUN-1997
CLASSIFICATION: 514

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IB95/00996
FILING DATE: 13-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/365,486
FILING DATE: 23-DEC-1994

ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 8255-0018.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 691 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO

ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: human enolase gene (EMBL #X56832)

INDIVIDUAL ISOLATE: fragment containing nucleotides -628 to +63
FEATURE:
NAME/KEY: CDS
LOCATION: 629..691

US-08-880-342-29

Query Match 52.0%; Score 13; DB 4; Length 691;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AGGAACCAAGAAAT 19
| | | | | | | | | |
Db 581 AGGAACCAAGAAAT 569

RESULT 25

US-08-991-789A-174
; Sequence 174, Application US/08991789A
; Patent No. 6225054
; GENERAL INFORMATION:

APPLICANT: Frudakis, Tony N.
APPLICANT: Smith, John M.
APPLICANT: Reed, Steven G.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF BREAST CANCER
NUMBER OF SEQUENCES: 292
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed IP Law Group
STREET: 701 Fifth Avenue, Suite 6300

```
;
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,789A
; FILING DATE: 11-Dec-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 210121.419C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
;
; INFORMATION FOR SEQ ID NO: 174:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 700 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 174:
US-08-991-789A-174

Query Match          52.0%; Score 13; DB 4; Length 700;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TGCAGGACCAAC 15
Db 203 TGCAGGACCAAC 215

RESULT 26
US-09-062-451-174
; Sequence 174, Application US/09062451
; Patent No. 6344550
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 297
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/062,451
; FILING DATE: 04-APR-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.419C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 174:
```

```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 700 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; US-09-062-451-174
;
; Query Match          52.0%; Score 13; DB 4; Length 700;
; Best Local Similarity 100.0%; Pred. No. 68;
; Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 3 TGCAGGACCAAC 15
; Db 203 TGCAGGACCAAC 215
;
; RESULT 27
; US-09-598-326-174
; Sequence 174, Application US/09598326
; Patent No. 6423496
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 247
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group PLLC
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/598,326
; FILING DATE: 20-Jun-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 210121.419D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 174:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 700 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 174:
US-09-598-326-174

Query Match          52.0%; Score 13; DB 4; Length 700;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TGCAGGACCAAC 15
Db 203 TGCAGGACCAAC 215

RESULT 28
US-08-961-083-133
; Sequence 133, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
```

APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 133:
SEQUENCE CHARACTERISTICS:
LENGTH: 829 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-083-133

Query Match 52.0%; Score 13; DB 3; Length 829;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 GAACAGAAATTT 21
|||||
DB 47 GAACAGAAATTT 59

RESULT 29
US-08-554-612C-26/c
Sequence 26, Application US/08554612C
Patent No. 5747660
GENERAL INFORMATION:
APPLICANT: Orlicky, David
TITLE OF INVENTION: PROTAGLANDIN F2\ RECEPTOR REGULATORY
TITLE OF INVENTION: PROTEIN AND THERAPEUTIC USES
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 2200 Sand Hill Road, Suite 100
CITY: Menlo Park
STATE: California
COUNTRY: U.S.A.
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/554,612C
FILING DATE: No. 5747660ember 6, 1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela

REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: 06519/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 322-5070
TELEFAX: (415) 854-0875
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 836 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-554-612C-26

Query Match 52.0%; Score 13; DB 1; Length 836;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TGCCAGGAACCAG 15
|||||
DB 660 TGCCAGGAACCAG 648

RESULT 30
US-08-554-612C-20/c
Sequence 20, Application US/08554612C
Patent No. 5747660
GENERAL INFORMATION:
APPLICANT: Orlicky, David
TITLE OF INVENTION: PROTAGLANDIN F2\ RECEPTOR REGULATORY
TITLE OF INVENTION: PROTEIN AND THERAPEUTIC USES
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 2200 Sand Hill Road, Suite 100
CITY: Menlo Park
STATE: California
COUNTRY: U.S.A.
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/554,612C
FILING DATE: No. 5747660ember 6, 1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: 06519/004001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 322-5070
TELEFAX: (415) 854-0875
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 932 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-554-612C-20

Query Match 52.0%; Score 13; DB 1; Length 932;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TGCCAGGAACCAG 15
|||||
DB 526 TGCCAGGAACCAG 514

```
RESULT 31
US-09-641-638-514/c
; Sequence 514, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GENSET.051CP1
; CURRENT APPLICATION NUMBER: US/09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 514
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 12-568-207 : polymorphic base G or T
; NAME/KEY: misc_binding
; LOCATION: 481..500
; OTHER INFORMATION: 12-568-207.mis1, potential
; NAME/KEY: misc_binding
; LOCATION: 502..521
; OTHER INFORMATION: 12-568-207.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 295..313
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 812..829
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc_binding
; LOCATION: 489..513
; OTHER INFORMATION: 12-568-207 potential probe
US-09-641-638-514

Query Match 52.0%; Score 13; DB 4; Length 1001;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AGGAACCGAGAAAT 19
Db 327 AGGAACCGAGAAAT 315

RESULT 32
US-09-641-638-515/c
; Sequence 515, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GENSET.051CP1
; CURRENT APPLICATION NUMBER: US/09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
```

```
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 515
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 12-568-365 : polymorphic base G or T
; NAME/KEY: misc_binding
; LOCATION: 481..500
; OTHER INFORMATION: 12-568-365.mis1, potential
; NAME/KEY: misc_binding
; LOCATION: 502..521
; OTHER INFORMATION: 12-568-365.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 137..155
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 654..671
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc_binding
; LOCATION: 489..513
; OTHER INFORMATION: 12-568-365 potential probe
US-09-641-638-515

Query Match 52.0%; Score 13; DB 4; Length 1001;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AGGAACCGAGAAAT 19
Db 169 AGGAACCGAGAAAT 157

RESULT 33
US-09-641-638-516/c
; Sequence 516, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GENSET.051CP1
; CURRENT APPLICATION NUMBER: US/09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 516
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
```


/ LOCATION: 501
/ OTHER INFORMATION: 12-568-367 : polymorphic base G or T
/ NAME/KEY: misc.binding
/ LOCATION: 481..500
/ OTHER INFORMATION: 12-568-367.mis1, potential
/ NAME/KEY: misc.binding
/ LOCATION: 502..521
/ OTHER INFORMATION: 12-568-367.mis2, potential complement
/ NAME/KEY: primer_bind
/ LOCATION: 135..153
/ OTHER INFORMATION: upstream amplification primer
/ NAME/KEY: primer_bind
/ LOCATION: 652..669
/ OTHER INFORMATION: downstream amplification primer, complement
/ NAME/KEY: misc.binding
/ LOCATION: 489..513
/ OTHER INFORMATION: 12-568-367 potential probe
US-09-641-638-516

Query Match 52.0%; Score 13; DB 4; Length 1001;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AGGAACACAGAAAT 19
Db 167 AGGAACACAGAAAT 155
|||||

RESULT 34
US-08-554-612C-17/c
; Sequence 17, Application US/08554612C
; Patent No. 5747660
; GENERAL INFORMATION:
; APPLICANT: Orlicky, David
; TITLE OF INVENTION: PROSTAGLANDIN F2' RECEPTOR REGULATORY
; TITLE OF INVENTION: PROTEIN AND THERAPEUTIC USES
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94025

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/554,612C
; FILING DATE: No. 5747660ember 6, 1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: 06519/004001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 322-5070
; TELEFAX: (415) 854-0875
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1007 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

Query Match 52.0%; Score 13; DB 1; Length 1007;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TGCCAGGAACCCAG 15
Db 526 TGCCAGGAACCCAG 514
|||||

RESULT 35
US-08-554-612C-18/c
; Sequence 18, Application US/08554612C
; Patent No. 5747660
; GENERAL INFORMATION:
; APPLICANT: Orlicky, David
; TITLE OF INVENTION: PROSTAGLANDIN F2' RECEPTOR REGULATORY
; TITLE OF INVENTION: PROTEIN AND THERAPEUTIC USES
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94025

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/554,612C
; FILING DATE: No. 5747660ember 6, 1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: 06519/004001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 322-5070
; TELEFAX: (415) 854-0875
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1007 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

US-08-554-612C-18

Query Match 52.0%; Score 13; DB 1; Length 1007;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TGCCAGGAACCCAG 15
Db 526 TGCCAGGAACCCAG 514
|||||

RESULT 36
US-09-276-531-129/c
; Sequence 129, Application US/09276531
; Patent No. 6183968
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Reddy, Roopa
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Marian R.
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
; TITLE OF INVENTION: RECEPTORS AND PROTEINS ASSOCIATED WITH CELL PROLIFERATION
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/276,531
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/079,677
; FILING DATE: March 27, 1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lynn E. Murry, Ph.D.
; REGISTRATION NUMBER: 42,918
; REFERENCE/DOCKET NUMBER: PA-0008 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 129:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1356 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: HNT2AZS07
; CLONE: 3143858
; US-09-276-531-129

Query Match 52.0%; Score 13; DB 4; Length 1356;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TGCAGGAACAC 15
|||||

DB 465 TGCAGGAACAC 453

RESULT 37
US-08-554-612C-16/c
; Sequence 16, Application US/08554612C
; Patent No. 5747660
; GENERAL INFORMATION:
; APPLICANT: Orlicky, David
; TITLE OF INVENTION: PROSTAGLANDIN F2₃ RECEPTOR REGULATORY
; TITLE OF INVENTION: PROTEIN AND THERAPEUTIC USES
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/554,612C
; FILING DATE: No. 5747660ender 6, 1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: 06519/004001

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 322-5070
; TELEFAX: (415) 854-0875
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1508 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-08-554-612C-16

Query Match 52.0%; Score 13; DB 1; Length 1508;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TGCAGGAACAC 15
|||||

DB 1027 TGCAGGAACAC 1015

RESULT 38
US-08-599-968-2/c
; Sequence 2, Application US/08599968
; Patent No. 6084153
; GENERAL INFORMATION:
; APPLICANT: Good, Allen G.
; APPLICANT: Stroeher, Virginia L.
; APPLICANT: Muench, Douglas G.
; TITLE OF INVENTION: Plants Having Enhanced Nitrogen
; TITLE OF INVENTION: Assimilation/Metabolism
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BENNETT JONES VERCHERE
; STREET: Suite 1800, 350 Albert Street
; CITY: Ottawa
; STATE: Ontario
; COUNTRY: Canada
; ZIP: K1R 1A4
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO),
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/599,968
; FILING DATE: February 14, 1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: CALDWELL, Roseann B.
; REGISTRATION NUMBER: 37077
; REFERENCE/DOCKET NUMBER: 10230-6
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1701 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-599-968-2

Query Match 52.0%; Score 13; DB 3; Length 1701;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGAACC 13
|||||

DB 1438 TGTGCCAGGAACC 1426

RESULT 39
US-08-487-596-9

```
; Sequence 9, Application US/08487596
; Patent No. 6440681
; GENERAL INFORMATION:
; APPLICANT: Elliot, Kathryn J.
; APPLICANT: Ellis, Steven B.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING AGONISTS AND
; APPLICANT: Harpold, Michael M.
; TITLE OF INVENTION: ANTAGONISTS FOR HUMAN NEURONAL
; TITLE OF INVENTION: NICOTINIC ACETYLCHOLINE RECEPTORS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,596
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO US94/02447
; FILING DATE: 08-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/149,503
; FILING DATE: 08-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/028,031
; FILING DATE: 08-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/938,154
; FILING DATE: 30-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/504,455
; FILING DATE: 03-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-9951
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0959
; TELEFAX: 619-238-0062
; INFORMATION FOR SEQ ID NO: 9:
; LENGTH: 1743 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 143..1627
; OTHER INFORMATION: /product= "ALPHA-6 SUBUNIT"
; US-08-487-596-9

Query Match 52.0%; Score 13; DB 4; Length 1743;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 CAGAAATTACAG 25
| | | | | | | | | |
Db 1648 CAGAAATTACAG 1660

RESULT 40
PCT-US92-01015-1
; Sequence 1, Application PC/TUS9201015
```

```
; GENERAL INFORMATION:
; APPLICANT: Davis, Geneva R
; APPLICANT: Provow, Sally P
; TITLE OF INVENTION: Production of Human Serum Albumin in
; TITLE OF INVENTION: Methylophic Yeast Cells
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fitch, Even, Tabin & Flannery
; STREET: 135 South LaSalle Street, Suite 900
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/01015
; FILING DATE: 19920204
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/650,040
; FILING DATE: 04-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 50857PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)552-1311
; TELEFAX: (619)552-0095
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1758 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1758
; OTHER INFORMATION:
; PCT-US92-01015-1

Query Match 52.0%; Score 13; DB 5; Length 1758;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CAGGAACCAAGAA 18
| | | | | | | | | |
Db 280 CAGGAACCAAGAA 292

RESULT 41
US-09-149-476-120
; Sequence 120, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
```

1	EARLIER FILING DATE: 1997-03-07
2	EARLIER APPLICATION NUMBER: 60/040,334
3	EARLIER FILING DATE: 1997-03-07
4	EARLIER APPLICATION NUMBER: 60/040,336
5	EARLIER FILING DATE: 1997-03-07
6	EARLIER APPLICATION NUMBER: 60/040,163
7	EARLIER FILING DATE: 1997-03-07
8	EARLIER APPLICATION NUMBER: 60/047,600
9	EARLIER FILING DATE: 1997-05-23
10	EARLIER APPLICATION NUMBER: 60/047,615
11	EARLIER FILING DATE: 1997-05-23
12	EARLIER APPLICATION NUMBER: 60/047,597
13	EARLIER FILING DATE: 1997-05-23
14	EARLIER APPLICATION NUMBER: 60/047,502
15	EARLIER FILING DATE: 1997-05-23
16	EARLIER APPLICATION NUMBER: 60/047,633
17	EARLIER FILING DATE: 1997-05-23
18	EARLIER APPLICATION NUMBER: 60/047,583
19	EARLIER FILING DATE: 1997-05-23
20	EARLIER APPLICATION NUMBER: 60/047,617
21	EARLIER FILING DATE: 1997-05-23
22	EARLIER APPLICATION NUMBER: 60/047,618
23	EARLIER FILING DATE: 1997-05-23
24	EARLIER APPLICATION NUMBER: 60/047,503
25	EARLIER FILING DATE: 1997-05-23
26	EARLIER APPLICATION NUMBER: 60/047,592
27	EARLIER FILING DATE: 1997-05-23
28	EARLIER APPLICATION NUMBER: 60/047,581
29	EARLIER FILING DATE: 1997-05-23
30	EARLIER APPLICATION NUMBER: 60/047,584
31	EARLIER FILING DATE: 1997-05-23
32	EARLIER APPLICATION NUMBER: 60/047,500
33	EARLIER FILING DATE: 1997-05-23
34	EARLIER APPLICATION NUMBER: 60/047,587
35	EARLIER FILING DATE: 1997-05-23
36	EARLIER APPLICATION NUMBER: 60/047,492
37	EARLIER FILING DATE: 1997-05-23
38	EARLIER APPLICATION NUMBER: 60/047,598
39	EARLIER FILING DATE: 1997-05-23
40	EARLIER APPLICATION NUMBER: 60/047,613
41	EARLIER FILING DATE: 1997-05-23
42	EARLIER APPLICATION NUMBER: 60/047,582
43	EARLIER FILING DATE: 1997-05-23
44	EARLIER APPLICATION NUMBER: 60/047,596
45	EARLIER FILING DATE: 1997-05-23
46	EARLIER APPLICATION NUMBER: 60/047,612
47	EARLIER FILING DATE: 1997-05-23
48	EARLIER APPLICATION NUMBER: 60/047,632
49	EARLIER FILING DATE: 1997-05-23
50	EARLIER APPLICATION NUMBER: 60/047,601
51	EARLIER FILING DATE: 1997-05-23
52	EARLIER APPLICATION NUMBER: 60/043,580
53	EARLIER FILING DATE: 1997-04-11
54	EARLIER APPLICATION NUMBER: 60/043,568
55	EARLIER FILING DATE: 1997-04-11
56	EARLIER APPLICATION NUMBER: 60/043,314
57	EARLIER FILING DATE: 1997-04-11
58	EARLIER APPLICATION NUMBER: 60/043,569
59	EARLIER FILING DATE: 1997-04-11
60	EARLIER APPLICATION NUMBER: 60/043,311
61	EARLIER FILING DATE: 1997-04-11
62	EARLIER APPLICATION NUMBER: 60/043,671
63	EARLIER FILING DATE: 1997-04-11
64	EARLIER APPLICATION NUMBER: 60/043,674
65	EARLIER FILING DATE: 1997-04-11
66	EARLIER APPLICATION NUMBER: 60/043,669
67	EARLIER FILING DATE: 1997-04-11
68	EARLIER APPLICATION NUMBER: 60/043,312
69	EARLIER FILING DATE: 1997-04-11
70	EARLIER APPLICATION NUMBER: 60/043,313
71	EARLIER FILING DATE: 1997-04-11
72	EARLIER APPLICATION NUMBER: 60/043,672
73	EARLIER FILING DATE: 1997-04-11

1	EARLIER APPLICATION NUMBER: 60/043,315
2	EARLIER FILING DATE: 1997-04-11
3	EARLIER APPLICATION NUMBER: 60/048,974
4	EARLIER FILING DATE: 1997-06-06
5	EARLIER APPLICATION NUMBER: 60/056,886
6	EARLIER FILING DATE: 1997-08-22
7	EARLIER APPLICATION NUMBER: 60/056,877
8	EARLIER FILING DATE: 1997-08-22
9	EARLIER APPLICATION NUMBER: 60/056,889
10	EARLIER FILING DATE: 1997-08-22
11	EARLIER APPLICATION NUMBER: 60/056,893
12	EARLIER FILING DATE: 1997-08-22
13	EARLIER APPLICATION NUMBER: 60/056,630
14	EARLIER FILING DATE: 1997-08-22
15	EARLIER APPLICATION NUMBER: 60/056,878
16	EARLIER FILING DATE: 1997-08-22
17	EARLIER APPLICATION NUMBER: 60/056,662
18	EARLIER FILING DATE: 1997-08-22
19	EARLIER APPLICATION NUMBER: 60/056,879
20	EARLIER FILING DATE: 1997-08-22
21	EARLIER APPLICATION NUMBER: 60/056,882
22	EARLIER FILING DATE: 1997-08-22
23	EARLIER APPLICATION NUMBER: 60/056,637
24	EARLIER FILING DATE: 1997-08-22
25	EARLIER APPLICATION NUMBER: 60/056,903
26	EARLIER FILING DATE: 1997-08-22
27	EARLIER APPLICATION NUMBER: 60/056,888
28	EARLIER FILING DATE: 1997-08-22
29	EARLIER APPLICATION NUMBER: 60/056,911
30	EARLIER FILING DATE: 1997-08-22
31	EARLIER APPLICATION NUMBER: 60/056,636
32	EARLIER FILING DATE: 1997-08-22
33	EARLIER APPLICATION NUMBER: 60/056,874
34	EARLIER FILING DATE: 1997-08-22
35	EARLIER APPLICATION NUMBER: 60/056,910
36	EARLIER FILING DATE: 1997-08-22
37	EARLIER APPLICATION NUMBER: 60/056,864
38	EARLIER FILING DATE: 1997-08-22
39	EARLIER APPLICATION NUMBER: 60/056,631
40	EARLIER FILING DATE: 1997-08-22
41	EARLIER APPLICATION NUMBER: 60/056,845
42	EARLIER FILING DATE: 1997-08-22
43	EARLIER APPLICATION NUMBER: 60/056,892
44	EARLIER FILING DATE: 1997-08-22
45	EARLIER APPLICATION NUMBER: 60/057,761
46	EARLIER FILING DATE: 1997-08-22
47	EARLIER APPLICATION NUMBER: 60/047,595
48	EARLIER FILING DATE: 1997-05-23
49	EARLIER APPLICATION NUMBER: 60/047,599
50	EARLIER FILING DATE: 1997-05-23
51	EARLIER APPLICATION NUMBER: 60/047,588
52	EARLIER FILING DATE: 1997-05-23
53	EARLIER APPLICATION NUMBER: 60/047,586
54	EARLIER FILING DATE: 1997-05-23
55	EARLIER APPLICATION NUMBER: 60/047,590
56	EARLIER FILING DATE: 1997-05-23
57	EARLIER APPLICATION NUMBER: 60/047,594
58	EARLIER FILING DATE: 1997-05-23
59	EARLIER APPLICATION NUMBER: 60/047,589
60	EARLIER FILING DATE: 1997-05-23
61	EARLIER APPLICATION NUMBER: 60/047,593
62	EARLIER FILING DATE: 1997-05-23
63	EARLIER APPLICATION NUMBER: 60/047,614
64	EARLIER FILING DATE: 1997-05-23
65	EARLIER APPLICATION NUMBER: 60/043,578

; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match 52.0%; Score 13; DB 4; Length 1782;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CAGGAACACGAAA 18
| | | | | | | | | |
Db 707 CAGGAACACGAAA 719

RESULT 42
US-09-017-706-3/c
; Sequence 3, Application US/09017706A
; Patent No. 6087147
; GENERAL INFORMATION:
; APPLICANT: ITO, YOSHIFUMI
; TITLE OF INVENTION: A-AMYLASE GENE HAVING ABILITY FOR HIGHLY PRODUCING
; TITLE OF INVENTION: MALTOPEPTIDASE, VECTOR CONTAINING SAID GENE AND
; FILE REFERENCE: 8361-0003-0
; CURRENT APPLICATION NUMBER: US/09/017,706A
; EARLIER FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: JP 305071/1997
; EARLIER FILING DATE: 1997-10-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1981
; TYPE: DNA
; ORGANISM: Pseudomonas sp., Strain KO-8940
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1848)
; NAME/KEY: mat_peptide
; LOCATION: (85)..(1848)
; NAME/KEY: mat_peptide
; LOCATION: (85)..(1848)

; OTHER INFORMATION: METHOD OF DETERMINING CHARACTERISTIC: E
; FEATURE:
; OTHER INFORMATION: MUTATED GENOMIC DNA
; FEATURE:
; OTHER INFORMATION: PLASMID POS3410F57
US-09-017-706-3

Query Match 52.0%; Score 13; DB 3; Length 1981;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CCAGGAACACGAAA 17
| | | | | | | | | |
Db 730 CCAGGAACACGAAA 718

RESULT 43
US-09-017-706-4/c
; Sequence 4, Application US/09017706A
; Patent No. 6087147
; GENERAL INFORMATION:
; APPLICANT: ITO, YOSHIFUMI
; TITLE OF INVENTION: A-AMYLASE GENE HAVING ABILITY FOR HIGHLY PRODUCING
; TITLE OF INVENTION: MALTOPEPTIDASE, VECTOR CONTAINING SAID GENE AND
; FILE REFERENCE: 8361-0003-0
; CURRENT APPLICATION NUMBER: US/09/017,706A
; EARLIER FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: JP 305071/1997
; EARLIER FILING DATE: 1997-10-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1981
; TYPE: DNA
; ORGANISM: Pseudomonas sp., Strain KO-8940
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1848)
; NAME/KEY: mat_peptide
; LOCATION: (85)..(1848)
; NAME/KEY: mat_peptide
; LOCATION: (85)..(1848)
; OTHER INFORMATION: METHOD OF DETERMINING CHARACTERISTIC: E
; FEATURE:
; OTHER INFORMATION: MUTATED GENOMIC DNA
; FEATURE:
; OTHER INFORMATION: PLASMID POS3410H57
US-09-017-706-4

Query Match 52.0%; Score 13; DB 3; Length 1981;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 CCAGGAACACGAAA 17
| | | | | | | | | |
Db 730 CCAGGAACACGAAA 718

RESULT 44
US-09-017-706-5/c
; Sequence 5, Application US/09017706A
; Patent No. 6087147
; GENERAL INFORMATION:
; APPLICANT: ITO, YOSHIFUMI
; TITLE OF INVENTION: A-AMYLASE GENE HAVING ABILITY FOR HIGHLY PRODUCING
; TITLE OF INVENTION: MALTOPEPTIDASE, VECTOR CONTAINING SAID GENE AND
; FILE REFERENCE: 8361-0003-0
; CURRENT APPLICATION NUMBER: US/09/017,706A
; EARLIER FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: JP 305071/1997
; EARLIER FILING DATE: 1997-10-21
; NUMBER OF SEQ ID NOS: 14

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1981
; TYPE: DNA
; ORGANISM: Pseudomonas sp., Strain KO-8940
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1848)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (85)..(1848)
; OTHER INFORMATION: METHOD OF DETERMINING CHARACTERISTIC: E
; FEATURE:
; OTHER INFORMATION: MUTATED GENOMIC DNA
; FEATURE:
; OTHER INFORMATION: PLASMID pos34100L57
US-09-017-706-5
```

```
Query Match          52.0%; Score 13; DB 3; Length 1981;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      5 CCAGGAACCGAGAA 17
        |||||
Db      730 CCAGGAACCGAGAA 718
```

```
RESULT 45
US-09-017-706-6/c
; Sequence 6, Application US/09017706A
; Patent No. 6087147
; GENERAL INFORMATION:
; APPLICANT: ITO, YOSHIFUMI
; TITLE OF INVENTION: A-AMYLASE GENE HAVING ABILITY FOR HIGHLY PRODUCING
; TITLE OF INVENTION: MALTOPENTAOSE, VECTOR CONTAINING SAID GENE AND
; FILE REFERENCE: 8361-0003-0
; CURRENT APPLICATION NUMBER: US/09/017,706A
; EARLIER FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: JP 305071/1997
; EARLIER FILING DATE: 1997-10-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1981
; TYPE: DNA
; ORGANISM: Pseudomonas sp., Strain KO-8940
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1848)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (85)..(1848)
; FEATURE:
; OTHER INFORMATION: MUTATED GENOMIC DNA
; FEATURE:
; OTHER INFORMATION: PLASMID: pos3410F139
US-09-017-706-6
```

```
Query Match          52.0%; Score 13; DB 3; Length 1981;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      5 CCAGGAACCGAGAA 17
        |||||
Db      730 CCAGGAACCGAGAA 718
```

Search completed: November 24, 2002, 15:04:07
Job time : 71.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2002, 14:30:54 ; Search time 52.5 Seconds
(without alignments)
180.348 Million cell updates/sec

Title: US-09-784-423-125

Perfect score: 25

Sequence: 1 TGTGCGAGGACCAAGAAATTACAG 25

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 335578 seqs, 18935133 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671156

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published.Applications_NA.*

- 1: /cgn2_6/ptodata/1/pubnpa/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubnpa/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubnpa/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubnpa/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubnpa/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubnpa/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubnpa/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubnpa/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubnpa/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/1/pubnpa/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubnpa/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/1/pubnpa/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/1/pubnpa/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/1/pubnpa/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	25	100.0	25	10	US-09-784-423-125
C 2	25	100.0	1000	10	Sequence 125, Appl
C 3	16	64.0	442	10	Sequence 32, Appl
4	15	60.0	569	10	Sequence 11479, A
5	15	60.0	1302	10	Sequence 198, Appl
6	15	60.0	11821	10	Sequence 9647, Ap
7	15	60.0	15535	10	Sequence 2857, Ap
C 8	15	60.0	106323	10	Sequence 2855, Ap
9	14	56.0	193	10	Sequence 3, Appl
C 10	14	56.0	963	10	Sequence 430, Appl
11	14	56.0	2273	10	Sequence 1948, Ap
12	14	56.0	9493	8	Sequence 110, Appl
C 13	14	56.0	12989	10	Sequence 163, Appl
14	14	56.0	21833	10	Sequence 1489, Ap
15	14	56.0	137997	10	Sequence 2275, Ap
16	14	56.0	684973	10	Sequence 3, Appl
C 17	14	56.0	1503841	9	Sequence 1, Appl
C 18	14	56.0	1503841	10	Sequence 1, Appl
C 19	14	56.0	1503841	10	Sequence 1, Appl

Sequence 6, Appl
Sequence 1310, Ap
Sequence 1310, Ap
Sequence 1310, Ap
Sequence 11569, A
Sequence 2301, Ap
Sequence 14842, A
Sequence 6327, Ap
Sequence 11959, A
Sequence 10539, A
Sequence 117, App
Sequence 11124, A
Sequence 2711, Ap
Sequence 4653, Ap
Sequence 7956, Ap
Sequence 6412, Ap
Sequence 2744, Ap
Sequence 6944, Ap
Sequence 4307, Ap
Sequence 6095, Ap
Sequence 11922, A
Sequence 163, App
Sequence 5680, Ap
Sequence 7386, Ap
Sequence 119, App
Sequence 60, Appl

Sequence 6, Appl
Sequence 1310, Ap
Sequence 1310, Ap
Sequence 1310, Ap
Sequence 11569, A
Sequence 2301, Ap
Sequence 14842, A
Sequence 6327, Ap
Sequence 11959, A
Sequence 10539, A
Sequence 117, App
Sequence 11124, A
Sequence 2711, Ap
Sequence 4653, Ap
Sequence 7956, Ap
Sequence 6412, Ap
Sequence 2744, Ap
Sequence 6944, Ap
Sequence 4307, Ap
Sequence 6095, Ap
Sequence 11922, A
Sequence 163, App
Sequence 5680, Ap
Sequence 7386, Ap
Sequence 119, App
Sequence 60, Appl

ALIGNMENTS

RESULT 1

US-09-784-423-125

; Sequence 125, Application US/09784423

; Patent No. US20020012924A1

; GENERAL INFORMATION:

; APPLICANT: Schumm, James W.

; TITLE OF INVENTION: MATERIALS AND METHODS FOR

; IDENTIFYING AND ANALYZING INTERMEDIATE TANDEM

; REPEAT DNA MARKERS

; NUMBER OF SEQUENCES: 147

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Promega Corporation

; STREET: 2800 Woods Hollow Road

; CITY: Madison

; STATE: Wisconsin

; COUNTRY: U.S.A.

; ZIP: 53711-5399

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Mb

; COMPUTER: IBM compatible PC

; OPERATING SYSTEM: Windows 95

; SOFTWARE: Word 97 (DOS text format)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/784,423

; FILING DATE: 15-Feb-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/018,584

; FILING DATE: 04-Feb-1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Grady J. Frenchick

; REGISTRATION NUMBER: 29,018

; REFERENCE/DOCKET NUMBER: 16026.9180

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (608) 257-3501

; TELEFAX: (608) 257-2275

; INFORMATION FOR SEQ ID NO: 125

; SEQUENCE CHARACTERISTICS:

; LENGTH: 25

; TYPE: Nucleic Acid

; STRANDEDNESS: Single

TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 125
US-09-784-423-125

Query Match 100.0%; Score 25; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGACCAAGAAATTACAG 25
|||||
Db 1 TGTGCCAGGACCAAGAAATTACAG 25

RESULT 2

US-09-784-423-32/c

Sequence 32, Application US/09784423
Patent No. US20020012924A1

GENERAL INFORMATION:

APPLICANT: Schumm, James W.

INVENTOR: Bacher, Jeffery W.

TITLE OF INVENTION: MATERIALS AND METHODS FOR

IDENTIFYING AND ANALYZING INTERMEDIATE TANDEM

REPEAT DNA MARKERS

NUMBER OF SEQUENCES: 147

CORRESPONDENCE ADDRESS:

ADDRESSEE: Promega Corporation

STREET: 2800 Woods Hollow Road

CITY: Madison

STATE: Wisconsin

COUNTRY: U.S.A.

ZIP: 53711-5399

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Mb

COMPUTER: IBM compatible PC

OPERATING SYSTEM: Windows 95

SOFTWARE: Word 97 (DOS text format)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/784,423

FILING DATE: 15-Feb-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/018,584

FILING DATE: 04-Feb-1998

ATTORNEY/AGENT INFORMATION:

NAME: Grady J. Frenchick

REGISTRATION NUMBER: 29,018

REFERENCE/DOCKET NUMBER: 16026,9180

TELEPHONE: (608) 257-3501

TELEFAX: (608) 257-2275

INFORMATION FOR SEQ ID NO: 32

SEQUENCE CHARACTERISTICS:

LENGTH: 1000 bp

TYPE: Nucleic Acid

STRANDEDNESS: Double

TOPOLOGY: Circular

MOLECULE TYPE: Genomic DNA

HYPOTHETICAL: no

IMMEDIATE SOURCE:

CLONE: 5132

POSITION IN GENOME:

CHROMOSOME/SEGMENT: 22

SEQUENCE DESCRIPTION: SEQ ID NO: 32

US-09-784-423-32

Query Match 100.0%; Score 25; DB 10; Length 1000;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGACCAAGAAATTACAG 25
|||||
Db 726 TGTGCCAGGACCAAGAAATTACAG 702

RESULT 3

US-09-864-761-11479/c

Sequence 11479, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

INVENTOR: Hanzel, David R.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE REFERENCE: Aemica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

PRIOR FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263,6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/774,203

PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Annonax Sequence Listing Engine vers. 1.1

SEQ ID NO 11479

LENGTH: 442

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AC005961.1

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.99

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.91

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.8

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1

US-09-864-761-11479

Query Match 64.0%; Score 16; DB 10; Length 442;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGACCAAGAAATTACAG 16
|||||

Db 226 TGTGCCAGGAACAGCA 211

RESULT 4

US-09-917-800A-198
; Sequence 198, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 198
; LENGTH: 569
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 AA875126
US-09-917-800A-198

Query Match 60.0%; Score 15; DB 10; Length 569;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GCCAGGAACAGCA 18
|||||
Db 240 GCCAGGAACAGCA 254

RESULT 5

US-09-815-242-9647
; Sequence 9647, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9647
; LENGTH: 1302
; TYPE: DNA
; ORGANISM: Salmonella typhi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1302)
US-09-815-242-9647

Query Match 60.0%; Score 15; DB 10; Length 1302;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 GGAACCAAGAAATTTA 22
|||||
Db 735 GGAACCAAGAAATTTA 749

RESULT 6

US-09-764-877-2857
; Sequence 2857, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2857
; LENGTH: 11821
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2857

Query Match 60.0%; Score 15; DB 10; Length 11821;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGAACAGCA 15
|||||
Db 3694 TGTGCCAGGAACAGCA 3708

RESULT 7

US-09-764-877-2855
; Sequence 2855, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper

```
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2855
; LENGTH: 15535
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2855

Query Match      60.0%; Score 15; DB 10; Length 15535;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGAGACCAG 15
Db 3783 TGTGCCAGGAGACCAG 3797

RESULT 8
US-09-803-661-3/c
; Sequence 3; Application US/09803661
; Patent No. US20020132292A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001162
; CURRENT APPLICATION NUMBER: US/09/803,661
; CURRENT FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 106323
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(106323)
; OTHER INFORMATION: n = A,T,C or G
US-09-803-661-3

Query Match      60.0%; Score 15; DB 10; Length 106323;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGAGACCAG 15
Db 42335 TGTGCCAGGAGACCAG 42321

RESULT 9
US-09-783-590-430
; Sequence 430; Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16,2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 430
; LENGTH: 193
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (11)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (81)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (156)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (160)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (170)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-430
```

```
Query Match      56.0%; Score 14; DB 10; Length 193;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 8 GGAACCAAGAAATTT 21
Db 67 GGAACCAAGAAATTT 80
```

```
RESULT 10
US-09-833-381-1948/c
; Sequence 1948; Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NO. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1948
; LENGTH: 963
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(963)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1948
```

```
Query Match      56.0%; Score 14; DB 10; Length 963;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 6 CAGGAACCAAGAAAT 19
Db 22 CAGGAACCAAGAAAT 9
```

```
RESULT 11
US-09-853-386-110
; Sequence 110; Application US/09853386
; Patent No. US20020049151A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Evelyn
; APPLICANT: Bresnihan, Barry
; APPLICANT: Conneely, Orla
; APPLICANT: Fitzgerald, Oliver
```

;; TITLE OF INVENTION: Therapeutic Approaches to Diseases by Suppression of the NURR
;; TITLE OF INVENTION: Subfamily of Nuclear Transcription Factors
;; FILE REFERENCE: P01972US1
;; CURRENT APPLICATION NUMBER: US/09/853,386
;; CURRENT FILING DATE: 2001-05-11
;; PRIOR APPLICATION NUMBER: US 60/203645
;; PRIOR FILING DATE: 2000-09-12
;; NUMBER OF SEQ ID NOS: 153
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 110
;; LENGTH: 2273
;; TYPE: DNA
;; ORGANISM: mus musculus
US-09-853-386-110

Query Match 56.0%; Score 14; DB 10; Length 2273;
Best Local Similarity 100.0%; Pred. No. 34; Mismatches 0; Indels 0; Gaps 0;
Matches 14; Conservative 0;

QY 3 TGCCAGGAACACGA 16
|||||
Db 1998 TGCCAGGAACACGA 2011

RESULT 12
US-08-424-550B-163
;; Sequence 163, Application US/08424550B
;; Patent No. US20020119447A1
;; GENERAL INFORMATION:
;; APPLICANT: JOHN N. SIMONS
;; APPLICANT: TAMI J. PILOT-MATIAS
;; APPLICANT: GEORGE J. DAWSON
;; APPLICANT: GEORGE G. SCHLAUDER
;; APPLICANT: SURESH M. DESAI
;; APPLICANT: THOMAS P. LEARY
;; APPLICANT: ANTHONY SCOTT MUERHOFF
;; APPLICANT: JAMES C. ERKER
;; APPLICANT: SHERI L. BUIJK
;; APPLICANT: ISA K. MUSHAWAR
;; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
;; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
;; NUMBER OF SEQUENCES: 716
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
;; STREET: 100 ABBOTT PARK ROAD
;; CITY: ABBOTT PARK
;; STATE: IL
;; COUNTRY: USA
;; ZIP: 60064-3500
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/424,550B
;; FILING DATE:
;; CLASSIFICATION: 435435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: FOREMSKI, PRISCILLA E.
;; REGISTRATION NUMBER: 33,207
;; REFERENCE/DOCKET NUMBER: 5527.PC.01
;; TELEPHONE: 708-937-6365
;; TELEFAX: 708-938-2623
;; INFORMATION FOR SEQ ID NO: 163:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9493 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
US-08-424-550B-163

Query Match 56.0%; Score 14; DB 8; Length 9493;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGCCAGGAACCA 14
|||||
Db 4037 TGTGCCAGGAACCA 4050

RESULT 13
US-09-764-847-1489/c
;; Sequence 1489, Application US/09764847
;; Patent No. US20020132767A1
;; GENERAL INFORMATION:
;; APPLICANT: Rosen et al.
;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;; FILE REFERENCE: PC009
;; CURRENT APPLICATION NUMBER: US/09/764,847
;; CURRENT FILING DATE: 2001-01-17
;; Prior application data removed - consult PALM or file wrapper
;; NUMBER OF SEQ ID NOS: 2003
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 1489
;; LENGTH: 12989
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-764-847-1489

Query Match 56.0%; Score 14; DB 10; Length 12989;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AGGAACCAAGAAATT 20
|||||
Db 8716 AGGAACCAAGAAATT 8703

RESULT 14
US-09-764-877-2275
;; Sequence 2275, Application US/09764877
;; Patent No. US20020147140A1
;; GENERAL INFORMATION:
;; APPLICANT: Rosen et al.
;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;; FILE REFERENCE: PC005
;; CURRENT APPLICATION NUMBER: US/09/764,877
;; CURRENT FILING DATE: 2001-01-17
;; Prior application data removed - refer to PALM or file wrapper
;; NUMBER OF SEQ ID NOS: 4031
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 2275
;; LENGTH: 21833
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-764-877-2275

Query Match 56.0%; Score 14; DB 10; Length 21833;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TGCCAGGAACACGA 16
|||||
Db 17007 TGCCAGGAACACGA 17020

RESULT 15
US-09-822-246-3
;; Sequence 3, Application US/09822246
;; Patent No. US20020142383A1
;; GENERAL INFORMATION:
;; APPLICANT: MERKULOV, Gennady et al.
;; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
US-09-822-246-3

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
 ; FILE OF INVENTION: AND USES THEREOF
 ; FILE REFERENCE: CLO01149
 ; CURRENT APPLICATION NUMBER: US/09/822,246
 ; CURRENT FILING DATE: 2001-04-02
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 197997
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(197997)
 ; OTHER INFORMATION: n = A,T,C or G
 ; US-09-822-246-3

Query Match 56.0%; Score 14; DB 10; Length 197997;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AGGAACAGAAATT 20
 |||||
 DB 142769 AGGAACAGAAATT 142782

RESULT 16
 US-09-263-959-1
 ; Sequence 1, Application US/09263959
 ; Patent No. US20020150891A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hood, Leroy E.
 ; APPLICANT: Rowen, Lee
 ; APPLICANT: Koop, Ben F.
 ; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH UTILIZE
 ; NUMBER OF SEQUENCES: 1279
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Seed and Berry LLP
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue
 ; CITY: Seattle
 ; STATE: Washington
 ; COUNTRY: US
 ; ZIP: 98104-7092
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/263,959
 ; FILING DATE: 05-MAR-1999
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mcmasters, David D.
 ; REGISTRATION NUMBER: 33,963
 ; REFERENCE/DOCKET NUMBER: 920010.426C2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (206) 622-4900
 ; TELEFAX: (206) 682-6031
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 684973 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-09-263-959-1

Query Match 56.0%; Score 14; DB 10; Length 684973;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GCCAGGAACAGAA 17
 |||||

DB 300613 GCCAGGAACAGAA 300626

RESULT 17
 US-09-946-807-1/c
 ; Sequence 1, Application US/09946807
 ; Patent No. US20020165144A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Stefansson, Hreinn
 ; APPLICANT: Steinhorsdottir, Valgerdur
 ; APPLICANT: Guicher, Jeffrey R.
 ; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
 ; FILE REFERENCE: 2345.2004-001
 ; CURRENT APPLICATION NUMBER: US/09/946,807
 ; CURRENT FILING DATE: 2001-09-05
 ; PRIOR APPLICATION NUMBER: US/09/795,668
 ; PRIOR FILING DATE: 2001-02-28
 ; PRIOR APPLICATION NUMBER: US 09/515,716
 ; PRIOR FILING DATE: 2000-02-28
 ; NUMBER OF SEQ ID NOS: 1531
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1503841
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(1531)
 ; OTHER INFORMATION: r=g or a
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(1531)
 ; OTHER INFORMATION: y=t/u or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(1531)
 ; OTHER INFORMATION: m=a or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(1531)
 ; OTHER INFORMATION: k=g or t/u
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(1531)
 ; OTHER INFORMATION: s=g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(1531)
 ; OTHER INFORMATION: w=a or t/u
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(1531)
 ; OTHER INFORMATION: b=g or c or t/u
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(1531)
 ; OTHER INFORMATION: d=a or g or t/u
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(1531)
 ; OTHER INFORMATION: h=a or c or t/u
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(1531)
 ; OTHER INFORMATION: v=a or g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(1531)
 ; OTHER INFORMATION: n=a or g or c or t/u
 ; US-09-946-807-1

Query Match 56.0%; Score 14; DB 9; Length 1503841;
 Best Local Similarity 100.0%; Pred. No. 46;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CAGGAACAGAAAT 19
|||||

Db 816791 CAGGAACAGAAAT 816778

RESULT 18

US-09-795-668-1/c
; Sequence 1, Application US/09795668
; Patent No. US20020045577A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn
; APPLICANT: Steinthorsdottir, Valgerdur
; APPLICANT: Gulcher, Jeffrey R.
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345 2004-001
; CURRENT APPLICATION NUMBER: US/09/795,668
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/515,716
; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1503841
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: r-g or a
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: y-t/u or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: m-a or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: k-g or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: s-g or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: w-a or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: b-g or c or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: d-a or g or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: h-a or c or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: v-a or g or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: n-a or g or c or t/u

US-09-795-668-1

Query Match 56.0%; Score 14; DB 10; Length 1503841;
Best Local Similarity 100.0%; Pred. No. 46;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CAGGAACAGAAAT 19
|||||

Db 816791 CAGGAACAGAAAT 816778

RESULT 19

US-09-795-686-1/c
; Sequence 1, Application US/09795686
; Patent No. US20020094954A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn
; APPLICANT: Steinthorsdottir, Valgerdur
; APPLICANT: Gulcher, Jeffrey R.
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345 2005-001
; CURRENT APPLICATION NUMBER: US/09/795,686
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/515,715
; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1503841
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: r-g or a
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: y-t/u or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: m-a or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: k-g or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: s-g or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: w-a or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: b-g or c or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: d-a or g or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: h-a or c or t/u
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: v-a or g or c
; NAME/KEY: misc_feature
; LOCATION: (1)...(1531)
; OTHER INFORMATION: n-a or g or c or t/u

US-09-795-686-1

Query Match 56.0%; Score 14; DB 10; Length 1503841;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CAGGAACAGAAAT 19
|||||

Db 816791 CAGGAACAGAAAT 816778

RESULT 20

US-09-895-382-6/c
; Sequence 6, Application US/09895382
; Patent No. US20020137150A1
; GENERAL INFORMATION:
; APPLICANT: OHTAKI, HIROMI
; APPLICANT: NAKAMURA, JUN
; APPLICANT: IZUI, HIROSHI
; APPLICANT: NAKAMATSU, TSUYOSHI